

AIR FORCE



AD-A212 766

HUMAN

RESOURCES

DTIC
ELECTE
SEP 18 1989

2

AIR FORCE INSTRUCTOR EVALUATION ENHANCEMENT:
EFFECTIVE TEACHING BEHAVIORS
AND ASSESSMENT PROCEDURES

Gary D. Borich

Universal Energy Systems, Incorporated
Human Factors and Logistics Division
4401 Dayton-Xenia Road
Dayton, Ohio 45432-1894

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

August 1989

Interim Technical Paper for Period September 1987 - May 1988

Approved for public release; distribution is unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

89 9 15 011

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

Form Approved
OMB No. 0704-0188

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-88-55	
6a. NAME OF PERFORMING ORGANIZATION Universal Energy Systems, Incorporated	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Training Systems Division	
6c. ADDRESS (City, State, and ZIP Code) 4401 Dayton-Xenia Road Dayton, Ohio 45432-1894		7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory	8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F41689-86-D-0052	
8c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601		10. SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO. 62205F	PROJECT NO. 7734
		TASK NO. 13	WORK UNIT ACCESSION NO. 01
11. TITLE (Include Security Classification) Air Force Instructor Evaluation Enhancement: Effective Teaching Behaviors and Assessment Procedures			
12. PERSONAL AUTHOR(S) Borich, G.D.			
13a. TYPE OF REPORT Interim	13b. TIME COVERED FROM Sep 87 TO May 88	14. DATE OF REPORT (Year, Month, Day) August 1989	15. PAGE COUNT 118
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) instructor evaluation, teaching methods, teaching effectiveness, training. (SAC-17)		
FIELD 05 05	GROUP 08 06		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>This paper investigates the feasibility of implementing an evaluation and feedback system for enhancing the teaching effectiveness of Air Training Command (ATC) instructors. State-of-the-art indicators of teaching effectiveness are reviewed and compared to instructor evaluation procedures in the public school sector, and those used by the Air Force, Army, and Navy. In addition, several measurement and observation procedures are described and prioritized in terms of their suitability for providing feedback to ATC instructors. Finally, recommendations are made on implementing an enhanced feedback system for increasing ATC instructor performance which complement procedures currently used by ATC.</p>			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy J. Allin, Chief, STINFO Branch		22b. TELEPHONE (Include Area Code) (512) 536-3877	22c. OFFICE SYMBOL AFHRL/SCV

August 1989

**AIR FORCE INSTRUCTOR EVALUATION ENHANCEMENT:
EFFECTIVE TEACHING BEHAVIORS
AND ASSESSMENT PROCEDURES**

Gary D. Borich

**Universal Energy Systems, Incorporated
Human Factors and Logistics Division
4401 Dayton-Xenia Road
Dayton, Ohio 45432-1894**

**TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601**

**Reviewed and submitted for publication by
Nestor K. Ovalle, II, Lt Col, USAF
Chief, Training Assessment Branch**

Accession For	
NTIS	CRA&I
DTIC	TAB
Unannounced	
Justification	
By	
Distribution	
Availability Codes	
Dist	Avail and/or Special
A-1	

This publication is primarily a working paper. It is published solely to document work performed.



SUMMARY

This paper reviews the research literature on teaching effectiveness in the public school classroom and suggests ways in which this literature could be useful in enhancing current Air Training Command (ATC) instructor evaluation procedures. Specifically, this paper identifies key indicators of teaching effectiveness from the literature, compares these indicators with those currently used on ATC Form 281 and by the Army and Navy, identifies several observational and measurement formats suitable for feedback enhancement, and provides conclusions and recommendations concerning the feasibility of a new ATC feedback enhancement initiative.

PREFACE

This technical paper represents the initial feasibility phase in a possibly longer-term effort to develop a feedback enhancement procedure for Air Training Command (ATC) instructors that would complement current ATC Form 281 and related evaluation procedures. Should a decision favorable to developing an observation/measurement procedure for enhancing feedback to ATC instructors follow from this paper, the development and pilot-testing of a proposed feedback system would constitute additional phases of the longer-term activity.

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. STATE-OF-THE-ART INDICATORS OF TEACHING EFFECTIVENESS DERIVED FROM THE RESEARCH LITERATURE	1
Clarity	2
Variety	2
Task Orientation	3
Engagement in the Learning Process	4
Moderate to High Success Rate	4
Summary of Key Behaviors	5
Clarity in the Classroom	5
Primary	8
Secondary	9
Variety in the Classroom	12
Primary	12
Secondary	14
Task Orientation in the Classroom	17
Primary	17
Secondary	19
Engagement in the Learning Process	21
Primary	21
Secondary	25
Success Rate in the Classroom	29
Primary	29
Secondary	31
III. INSTRUCTOR EVALUATION PROCEDURES IN THE PUBLIC SCHOOL SECTOR	34
Competency Testing	35
Performance Appraisal	36
IV. DESCRIPTION OF INSTRUCTOR EVALUATION FORMS IN THE AIR FORCE, ARMY, AND NAVY AND HOW THEY ARE USED IN THEIR INSTRUCTIONAL ENVIRONMENTS	40
Instructor Evaluation in the Air Force	40
Comparison of Items on ATC Form 281 with Indicator of Teaching Effectiveness from the Research Literature	43
Instructor Evaluation in the Army and Navy and Some Comparisons with Form 281	48
V. SURVEY OF OBSERVATION AND MEASUREMENT FORMATS SUITABLE FOR FEEDBACK ENHANCEMENT	51
Time Interval Observation	51
Summated Ratings	55
Scalogram Analysis (the Guttman Scale)	56
Checklists	57
Time Interval Observation	58

Table of Contents (Concluded)

	Page
Rating Scales	58
Checklist	58
Guttman Scale	59
Some Advantages and Limitations of Time Interval Observation	60
Summated Ratings (Likert and Semantic Differential)	60
Some Advantages and Limitations of the Checklist	61
Some Advantages and Limitations of the Guttman Scale	62
VI. SUMMARY AND RECOMMENDATIONS	64
Summary	64
Recommendations	65
ATC Form 281	66
Supervisor Training Course	68
Evaluation of Laboratory- and Field-Based Instruction	69
Instructor Peer Evaluation	70
Student Evaluation of Instructors	70
Technical Training Instructor's Course	71
REFERENCES	72

LIST OF FIGURES

Figure	Page
1 Sign System	52
2 Category System for Recording Sequential Events	53
3 Rating System after a One-hour Interval of Observation	54

LIST OF TABLES

Table	Page
1 Effective and Ineffective Indicators for Clarity	6
2 Effective and Ineffective Indicators for Variety	13
3 Effective and Ineffective Indicators for Task Orientation	18
4 Some Direct Instruction Functions	22
5 Some Indirect Instruction Functions	23
6 Effective and Ineffective Indicators for Engaging Students in the Learning Process	24
7 Effective and Ineffective Indicators for Moderate to High Rates of Success	30
8 Methods of Evaluation Reported by the Nation's 100 Largest School Districts	37
9 Comparison of Key Indicators of Teaching Effectiveness from the Research Literature with Items on ATC Form 281	44
10 Measurement Systems by Possible Data Providers	59
11 Ranking of Each of Four Measurement System Against the Six Criteria	63

**AIR FORCE INSTRUCTOR EVALUATION ENHANCEMENT:
EFFECTIVE TEACHING BEHAVIORS AND ASSESSMENT PROCEDURES**

I. INTRODUCTION

The purpose of this technical paper is to investigate the feasibility of implementing an evaluation and feedback system for enhancing the teaching effectiveness of Air Training Command (ATC) instructors. Its primary purpose is to review state-of-the-art literature on indicators of teaching effectiveness, compare those indicators to those currently in use in the public school sector and on ATC Form 281, Instructor Evaluation Checklist, and to prioritize several measurement and/or observation procedures suitable for providing feedback to the ATC instructor. The work detailed in this paper is based on the assumption that the teaching effectiveness indicators and measurement procedures described herein would be used to complement and enhance the Form 281 currently in use by ATC.

Accordingly, this document has been organized into the following additional sections:

II. State-of-the-art indicators of teaching effectiveness derived from the research literature. This section reviews the research conducted in public school classrooms from 1960 to the present and indicates those key instructor behaviors and procedures which have been found to increase learners' performance on cognitive tests of academic achievement.

III. Instructor evaluation procedures in the public school sector. This section reviews some of the evaluation and feedback systems currently in use by school systems and State Departments of Education to evaluate and provide feedback to public school teachers. The teaching effectiveness indicators used in one of these systems are compared with those from the research literature.

IV. Description of instructor evaluation forms in the Air Force, Army, and Navy and how they are used in their instructor environments. This section summarizes the current evaluation procedure used by ATC and compares the teaching effectiveness indicators on Form 281 with those reported in the research literature and those currently in use by the Army and Navy.

V. Survey of observation and measurement formats for feedback enhancement. This section reviews several practical means by which the indicators of teaching effectiveness presented in Section II could be recorded and used in a professional development and/or self-improvement feedback system for ATC instructors. A prioritization of formats and individuals available to serve as data recorders/observers is provided based on estimates of acceptability to users, ease of implementation, cost, objectiveness, and increases in feedback, validity, and reliability.

VI. Summary and recommendations. This section summarizes the main findings of this paper and discusses the implications and possible impact of an enhanced feedback system for increasing ATC instructor performance and the impact of such a system in other sectors of the ATC, such as training in the ATC instructor's and supervisor's course, revision of ATC Form 281, and development of a peer evaluation form, a field/laboratory-based version of Form 281, and a new student evaluation form.

**II. STATE-OF-THE-ART INDICATORS OF TEACHING EFFECTIVENESS
DERIVED FROM THE RESEARCH LITERATURE**

From the research literature dating from 1960 to the present, approximately 10 teacher behaviors have shown promising empirical relationships to desirable student outcomes, primarily

measured as achievement on classroom and standardized tests of academic performance. These 10 behaviors have come from studies at both the elementary and secondary school levels that have used classroom interaction analysis (Evertson & Green, 1986; Rosenshine & Furst, 1973) to directly observe sequences of student-teacher behavior (dialogue) in the classroom which correlate with student performance on cognitive tests of academic achievement. Five behaviors resulting from these studies have strong research support while another five have promising support and also appear logically related to effective teaching. The first five behaviors will be identified as "key" behaviors, since for purposes of this review these will be considered essential for effective teaching. The second five will be identified as "catalytic" or "helping" behaviors which may occur in various mixtures to assist in implementing the key behaviors. In this section an overview of these five key behaviors will be provided first and, then, the subdimensions or individual components of each of the five will be described. The five key teaching behaviors for which empirical research studies conducted in public school classrooms have provided strong support are:

1. Clarity
2. Variety
3. Task orientation
4. Engagement in the learning process
5. Moderate to high rates of success

Clarity

This behavior refers to how clear and interpretable an instructor's presentation is to the class. For example, are the points the instructor is making understandable? Is the instructor able to explain concepts clearly such that the students are able to follow in a logical step-by-step order? And, is the instructor's oral delivery to the class clear, audible and intelligible, and free of any distracting mannerisms that could impair a student's understanding of the content. One result from the research on teacher clarity is that instructors have been found to vary considerably on this behavior. That is, not all instructors are able to communicate clearly and directly to their students without wandering, without speaking above the "heads" of students or without speech patterns (for example, speaking too fast) that impair the clarity of what is being presented.

The research has indicated that instructors who teach with high degrees of clarity generally spend less time having to go over material and generally have their questions answered correctly the first time through the material, allowing more time for instruction. Clarity is a complex behavior, since it is related to many other so-called cognitive behaviors, such as an instructor's organization of the content, familiarity with the lesson, and delivery strategies (e.g., whether an instructor uses the proper mix of discussion, question and answer, and lecture at the appropriate times). Nevertheless, both the cognitive clarity and oral clarity of an instructor's presentation have been found to vary substantially among teachers which, in turn, has produced significant differences in student performance on cognitive tests of achievement.

Variety

This behavior refers to variability or flexibility of delivery during the presentation of a lesson. It includes, for example, the planned mixing of different classroom behaviors, such as lecturing, group discussion, question and answers, and independent practice. For example, if, across several related lessons, an instructor's lesson format utilized these and other instructional techniques to weave a single, unified theme or achieve a predetermined behavioral objective, the instructor could be said to be using a high degree of variety. Research has

indicated that the use of variety in instructional materials and techniques, in the types of reinforcements used, and in the types of feedback given to students is related to increases in student achievement.

One of the most popular and effective ways of creating variety during instruction is to ask varying types of questions. Many different types of questions have been identified (Gall et al., 1978) and, when integrated into the pacing and sequencing of a lesson, have provided a means of creating meaningful variation within a lesson. An instructor's knowledge of question-asking behavior and the ability to discriminate and use the many different types of question formats (e.g., fact questions, process questions, convergent questions, divergent questions) in the process of delivering a lesson have been found to correlate with student achievement. These and other types of questioning formats have been the theme of numerous texts and training materials (e.g., Cunningham, 1971; Hunkins, 1976; Survey, 1974) and continue to be a popular focus of classroom research.

Another aspect of variety in teaching is the use of learning materials, equipment, displays, and space in the classroom. This includes the visual variety of the classroom and the extent to which this variety can actually encourage student involvement with lesson content. The display of graphic materials, use of audiovisual devices, display of maps and globes, and organization of different reference materials have been considered by students and instructors to contribute to instructional variety, which, in turn, has been found to influence student behavior during instruction and performance on end-of-unit tests. In some studies the amount of attention in a classroom was greater for classrooms which used greater amounts of variety in classroom activities and materials; other studies have noted changes related to student test performance.

Task Orientation

This behavior refers to the degree to which the instructor is achievement-oriented and has high expectations with respect to his or her students. Task-related aspects of this behavior are the amount of time the instructor spends lecturing and asking questions, encouraging students to inquire or think independently, and the amount of intellectual or cognitive emphasis provided by the instructor. These aspects of task orientation often manifest themselves in the degree of concern the instructor has that all relevant material gets covered and learned, as opposed to a preoccupation with procedural matters or an exclusive concern that the students simply enjoy the class. This behavior also can be found in the extent to which the instructor communicates high performance expectations to the class early in the course of instruction. Although researchers agree that the classrooms should provide a friendly and considerate context in which to learn, most agree that achievement has been higher in classrooms that espouse primarily a business-like atmosphere and high performance expectations. That is, classrooms in which instructor-student interactions focus more on intellectual content and performance than on process issues (such as how to use materials or classroom rules and procedures) are more likely to have higher rates of achievement than those that do not. It has also been reported that the instructors in content-oriented classrooms were more conversant with the type of content likely to appear on department and standardized achievement tests. An instructor with a high degree of task orientation would be one who is goal-oriented--that is, who knows what behavioral outcomes he or she wants to achieve in a given period of time, organizes instruction around these observable outcomes, and adheres steadfastly to them in the midst of other less instructionally relevant activities which may arise from time to time to create distractions. The systematic organization of class content, the use of this organization to present performance-oriented lessons using teaching strategies that emphasize high performance expectations all have been found to be important ingredients of an instructor's task orientation.

Engagement in the Learning Process

As noted above, an instructor who is task-oriented is likely to maximize the amount of time a student has to learn the material to be tested. However, a distinctly different concept than the efficient use of instructional time is the time students are actually engaged in learning the material. This has been called the engagement rate of students, or the on-task behavior of students.

Engagement rate is the percentage of time devoted to learning in which the student is actually on-task and engaged with the instructional materials and activities being presented. The key to understanding engagement rate is the awareness that while an instructor can be task-oriented, providing maximum content coverage, the students may not be engaged or on-task during some or all of this time. This disengagement can involve an emotional or mental detachment from the lesson—or both. Off-task or nonengaged behavior can take both obvious and nonobvious forms. When students are physically inattentive (e.g., looking out the window or preoccupied with off-task materials), they obviously are not engaged in instruction, however thoroughly it is being presented by the instructor. However, research has shown that students can also be nonengaged in far more subtle ways, as with a student who looks attentive but whose thoughts are "miles" away. Correcting this type of nonengagement may be a far more difficult task. Several research studies have contributed useful data for increasing learning time and, most importantly, the student engagement rate. From these data it was possible for Crawford et al. (1978) to identify behaviors that have potential for increasing learning time that results in greater amounts of on-task behavior. Some of their suggestions are the following:

1. Instructors should have a system of rules that communicate to students what is expected and how learning will proceed in their classrooms prior to the start of the instruction (Brophy & Evertson, 1976; Stallings & Kaskowitz, 1974).
2. Instructors should frequently move around the room to monitor students' work and communicate to the students an awareness of their progress (McDonald & Elias, 1976; Stallings & Kaskowitz, 1974).
3. When students work independently, instructors should ensure that the assignments are challenging but can be completed by each student working without instructor direction (McDonald & Elias, 1976; Stallings & Kaskowitz, 1974).
4. Instructors should minimize such activities as giving directions and organizing the class for instruction, by writing the daily schedule on the board when appropriate, ensuring that students know how to proceed and what to do without asking or taking class time (McDonald & Elias, 1976; Soar & Soar, 1973).
5. Instructors should make abundant use of textbooks, workbooks, and other paper-and-pencil activities which are at or slightly above a student's current level of functioning (Brophy & Evertson, 1976; McDonald & Elias, 1976; Stallings & Kaskowitz, 1974).
6. Instructors should avoid "timing errors"; that is, they should prevent inattentiveness by making speedy transitions from one activity to another (Brophy & Evertson, 1976).

Moderate to High Success Rate

A crucial aspect of research that has investigated task orientation and student engagement has been the level of difficulty at which material is presented. Level of difficulty in these studies was measured as the rate at which students could understand and correctly complete

exercises pertaining to the material being taught. One possible level of difficulty is that of "high success," in which the student understands the task and makes only occasional careless errors; another is "medium success," in which the student has partial understanding, but makes some substantive errors; and a third is "low success," in which the student does not understand the task at all. The findings for task orientation and student engagement reported above are highly related to level of difficulty--or success rate. The findings consistently point out that instruction that produces a moderate to high success rate results in increased achievement. In addition, research has shown that instruction producing low error rates can contribute to high levels of student self-esteem and to positive attitudes toward the subject matter. The average student in a typical public school classroom spends about half of the time working on tasks that provide the opportunity for high success. Researchers have found that students who spend more than half their time in high success activities had higher achievement scores, better retention, and more positive attitudes toward their class. These findings have led to at least one suggestion that students spend from 60-70% of their time on tasks that afford the opportunity for moderate to high levels of success (Brophy & Evertson, 1976); that is, tasks that allow for almost complete understanding and only occasional careless errors the first time through the material.

Moderate to high success rates tend to produce mastery of the lesson content which also increases the likelihood that the student will be able to apply learned knowledge at a later time in some practical way (e.g., solving problems, making decisions, choosing correct solutions). Exercises providing moderate to high success rates allow the individual elements or pieces that are learned to "fall into place," thereby providing responses that are smooth and automatic. Research has indicated that many instructors devote insufficient time to this stage of learning (Fisher et al., 1978) which is believed to be particularly crucial for mastery of the content and the generalization of the content learned to the on-the-job training environment. Organizing and planning instruction that yields moderate to high success rates but at the same time is not boring, repetitive or time-wasting are considered to be a key behavior for effective teaching.

Summary of Key Behaviors

The five concepts of clarity, variety, task orientation, student engagement rate, and success rate are considered to be among the most important general teaching behaviors comprising modern definitions of effective teaching. It is generally reported that without the knowledge and skill to present lessons that are clear, that incorporate variety, that are task-oriented, and that actually engage students in the learning process at moderate to high rates of success, an instructor would not be effective in producing desirable patterns of student performance. In the remaining portions of this section we will present the subcomponents of each of these five behaviors and identify some observable signs by which they can be measured and feedback provided to the ATC instructor. In addition, the following discussion will identify five helping or catalytic behaviors generally considered useful in implementing the five key behaviors. These helping behaviors are the instructor's use of structuring, questioning, probing, student ideas, and enthusiasm.

Clarity in the Classroom

Table 1 summarizes some of the subdimensions or specific components by which current research has defined the concept of instructional clarity. The first two columns in Table 1 list some indicators of clarity likely to be observed among effective and ineffective instructors. The third column suggests some of the types of feedback that might be given to an instructor who is observed exhibiting one or more ineffective indices of clarity.

Table 1. Effective and Ineffective Indicators for Clarity

Being clear (An effective instructor)	Poor clarity (An ineffective instructor)	(Recommended behavior)
1. Informs learners of the lesson objective; e.g., describes what behaviors will be tested or required on future assignments as a result of the lesson.	Fails to link lesson content to how and at what level of complexity the content will be used.	Use behavioral objectives at the desired level of complexity (e.g., knowledge, comprehension). Indicate to the learners at the start of the lesson in what ways the behavior will be used in the future.
2. Provides learners with an advanced organizer; e.g., places lesson in perspective of past and/or future lessons.	Starts presenting content without first introducing the subject with respect to some broader context.	Consult a unit plan to determine what task-relevant prior learning is required for this lesson and what task-relevant prior learning this lesson represents for future lessons. Begin the lesson by informing the learner that the content to be taught is part of this larger context.
3. Checks for task-relevant prior learning at beginning of the lesson; e.g., determines level of understanding of prerequisite facts or concepts and reteaches, if necessary.	Moves to new content without checking for the facts, concepts or skills needed to acquire the new learning.	Ask questions of students at beginning of lesson or check assignments regularly to determine if task-relevant prior knowledge has been acquired.
4. Gives directions slowly and distinctly; e.g., repeats directions when needed or divides them into smaller chunks.	Presents too much clerical, managerial or technical information at once, too quickly.	Organize procedures for lengthy assignments in step-by-step order and give as handout as well as orally.
5. Knows ability levels and teaches at or slightly above learners' current level functioning; e.g., knows students' attention spans.	Fails to know that instruction is under or over heads of students. Seems not to know when most learners have "tuned out."	Determine ability level from tests, previous assignments and interests and retarget instruction accordingly.

Table 1. (Concluded)

Being clear (An effective instructor)	Poor clarity (An ineffective instructor)	(Recommended behavior)
6. Uses examples, illustrations and demonstrations to explain and clarify; e.g., uses visuals to help interpret and reinforce main points.	Restricts presentation to routine verbal reproduction of text or workbook; i.e., fails to personalize lesson ideas.	Restate main points in at least one modality other than one in which they were initially taught (e.g., visual versus auditory).
7. Provides review or summary at end of each lesson.	Ends lesson abruptly without "repackaging" key points.	Use key abstractions, repetition or symbols to help learners efficiently store and later recall content.

The research evidence. The primary and secondary sources which have been consulted which provide strong or promising research evidence for the subdimensions of clarity presented in Table 1 are as follows:¹

Primary

Alexander, L., Frankiewuz, R., & Williams, R. (1979). Facilitation of learning and retention of oral instruction using advance and post organizers. Journal of Educational Psychology, 71, 701-707.

Ausubel, D.P. (1960). The use of advanced organizers in the learning and retention of meaningful material. Journal of Educational Psychology, 51, 267-272.

Blaney, R. (1983). Effects of teacher structuring and reacting on achievement. Elementary School Journal, 83, 569-577.

Brophy, J.E., & Evertson, C.M. (1976). Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.

Clark, C., Gage, N., Marx, R., Peterson, P., Strayrouk, N., & Winnie, P. (1979). A factorial experiment on teacher structuring, soliciting and reacting. Journal of Educational Psychology, 71, 534-552.

Evertson, C., Emmer, E.T., & Brophy, J.E. (1980). Predictors of effective teaching in junior high mathematics classrooms. Journal of Research in Mathematics Education, 11, 169-178.

Good, T.L., & Grouws, D.A. (1979). The Missouri mathematics effectiveness project. Journal of Educational Psychology, 71, 355-362.

Hiller, J.H., Fisher, G., & Kaess, W. (1969). A computer investigation of verbal characteristics of effective classroom lecturing. American Educational Research Journal, 6, 161-175.

Kennedy, J., Cruickshank, D., Bush, A., & Myers, B. (1978). Additional investigations into the nature of clarity. Journal of Educational Research, 72, 3-10.

Land, M. (1979). Low-inference variables and teacher clarity: Effects on student concept learning. Journal of Educational Psychology, 71, 795-799.

Land, M., & Smith, L. (1979). Effect of a teacher clarity variable on student achievement. Journal of Educational Research, 72, 196-197.

Land, M., & Smith, L. (1979). The effect of low inference teacher clarity inhibitors on student achievement. Journal of Teacher Education, 30, 55-57.

Smith, L., & Cotten, M. (1980). Effect of lesson vagueness and discontinuity on student achievement and attitudes. Journal of Educational Psychology, 72, 670-675.

Smith, L., & Land, M. (1981). Low inference verbal behaviors related to teacher clarity. Journal of Classroom Interaction, 17, 37-42.

¹"Primary" refers to an original research study. "Secondary" refers to a review of research or meta analysis which summarizes the results of more than a single study.

Smith, L., & Sanders, K. (1981). The effects on student achievement and student perception of varying structure in social studies content. Journal of Educational Research, 74, 333-336.

Secondary

Belgard, M., Rosenshine, B., & Gage, N.L. (1971). Effectiveness in explaining: Evidence on its generality and correlation with pupil ratings. In Ian Westbury & A.A. Bellack (Eds.), Research into classroom processes: Recent developments and next steps. New York: Teachers College Press.

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 328-374). New York: Macmillan.

Cohen, P.A., Ebeling, B.J., & Kulik, J.A. (1981). A meta-analysis of outcome studies of visual-based instruction. Education Communication and Technology Journal, 29, 26-36.

Dunkin, M.J., & Biddle, B.J. (1974). The study of teaching. New York: Holt.

Luiten, J., Ames, W., & Aerson (1980). A meta-analysis of advanced organizers on learning and retention. American Educational Research Journal, 17, 211-218.

Melton, R. (1978). Resolution of conflicting claims concerning the effect of behavioral objectives on student learning. Review of Educational Research, 48, 291-302.

Rosenshine, B. (1971). Teaching behaviors and student achievement. London: National Foundation for Educational Research in England and Wales.

Rosenshine, B. (1968). To explain: A review of research. Educational Leadership, 26, 275-280.

Rationale for the behaviors comprising teacher clarity. The following section presents an overview of the rationale for each of the subcomponents of clarity based on the above-cited research literature.

As is noted in Table 1, the ability to be "clear" involves seven important behaviors. Three of these pertain to behaviors that should occur at the beginning of a lesson, whereas the other four pertain to behaviors that are implemented during a lesson. The three behaviors that should occur at the beginning of a lesson are:

1. Informing the learner of the objective.
2. Providing learners with an advanced organizer.
3. Checking for task-relevant prior learning and reteaching, if necessary.

Being clear means being understood by learners. However, being understood is believed to depend as much on what an instructor does prior to teaching a lesson as what is done during it. Therefore, these three "pre-instructional" activities are recommended to establish a learning set which focuses learners in the intended direction.

Typically, at the start of a lesson, learners do not know what behaviors, skills or concepts will be expected of them. This is considered to be an unintentional source of anxiety, since in the absence of information, unrealistic expectations may be created in the minds of learners which are anxiety-producing. Anxiety, and especially high levels of anxiety, has been reported to shorten learners' attention spans and provide a timely reason for not "tuning-in" the instruction. Informing learners of the objective conveys the expectations the instructor has for

them and dispels any unrealistic fears and expectations that might have been created. If learners know what to expect and, if the expectation is reasonable in their eyes, a desirable learning set is established.

Advanced organizers at the start of the lesson are often believed necessary to keep learners tuned-in and on-task. Advanced organizers place the day's learning in perspective by connecting it with prior and future parts of the curriculum. With advanced organizers a lesson is more likely to be seen by learners as important, as part of something larger than the lesson itself and, therefore, as more valued as part of this larger context. If the content to be taught appears irrelevant or unattached to anything that has gone before or is likely to come after, it often becomes content for which learners believe they will not be held accountable.

A third pre-instructional activity for achieving clarity involves checking for task-relevant prior learning. Without the mastery of task-relevant prior learning, a lesson may be incomprehensible to some, if not most, learners. A lesson is thought to have clarity only to the extent that all or the majority of students can exhibit the prerequisite facts, skills or concepts required for understanding the lesson. To achieve this, inquiries are made before or at the beginning of a lesson to see if learners can actually exhibit the prerequisite behaviors for new learning to occur. Many options are suggested in the research literature for conducting such an inquiry, from informal oral questions asked randomly of a few students to a more formal checking of homework or assignments. When extensive deficiencies in task-relevant prior knowledge are found, reteaching old but prerequisite content is considered more important than teaching new content scheduled for the day's lesson. Some techniques commonly recommended for checking task-relevant prior learning include:

1. Preparing on index cards, before the class begins, a pool of questions which some students will be asked to randomly choose from and orally answer at the start of class.
2. Asking a single brief question on the main point taught in the previous lesson.
3. Calling on students to summarize one major idea of their own choosing from a previous, related lesson.
4. Asking students to apply a specific rule, regulation or process that was taught in the previous lesson.
5. Having students write out the answers to several key questions placed on the board, while the instructor circulates and checks answers.
6. Having students write the answers to several key questions and check them with a classmate.
7. Having students prepare questions about the previous lesson and ask them of each other.

The four behaviors that should occur during a lesson in order to achieve clarity are:

1. Giving directives slowly and distinctly.
2. Knowing the ability levels of the learners and teaching at or slightly above their current level of functioning.
3. Using examples, illustrations and demonstrations to explain and clarify text and workbook content.
4. Providing a review or summary at the end of each lesson.

These behaviors differ from the first three in that they pertain to the actual presentation of lesson content, although preparation for them may begin long before the start of a lesson. These four practices also represent ways of establishing a clear and understandable dialogue with students. To "understand" is to receive the message being sent. As was noted with the three behaviors above, part of successfully receiving the message is being ready and prepared to receive it. Informing the learner as to what is expected, providing advanced organization and checking for and reteaching task-relevant prior learning are considered ways of helping to prepare learners to receive the message. In addition to these behaviors, however, researchers have studied ways to make clear the message itself; that is, to "package" it in ways that make it intelligible to learners, so that they can not only hear and interpret but also retain what is being said.

A frequently reported criticism from students is that they "weren't told" or "didn't understand" what to do. In addition, researchers report that learners frequently are afraid to admit that they could not follow the directions given, either because they could not remember all that was said or it was said so quickly and matter-of-factly that the most important part of the directive became lost among less relevant detail. The point of this fourth behavior is directions pertaining to completing exercises, reading assignments, homework, and drill and practice need to be communicated with the same deliberateness as used to present content. Typically, it has been found that instructors speed up and speak less distinctly when giving directions as to what to do, how to proceed, and what rules to follow in completing work than when actually teaching the content to which the directions pertain. When learners cannot follow directions as to how the instructor wants them to become engaged in the learning process or complete performance-oriented work assignments, researchers report that more often than not the students silently proceed on their own, many times missing the intent of an assignment. The appropriate recommendation often given here is to slow down when conveying instructions, taking care to divide directives into steps if need be, and checking to be sure each step is understood along the way.

The next behavior, knowing the ability levels of the students and teaching at their levels, often is reported to be one of the most difficult to achieve. Since most instructors are likely to have students of several levels of ability (e.g., less able, average, more able) represented in their classes, they are likely to become frustrated by an inability to teach some students. Nevertheless, research on clarity suggests the need to package instructional stimuli in the form of oral presentations, visual messages, exercises and reading assignments that are at the current level of functioning of all of the students. To accomplish this, instructors may need to provide a range of instructional stimuli in their classrooms. Demonstration centers, reference libraries, different types of pictorial displays, and even alternate texts and exercises that tap into abilities slightly above and below the average learner are some of the ways researchers suggest that clarity can be maintained in a diverse classroom.

The next two behaviors for achieving clarity both occur during the lesson, and both are intended to expand upon and clarify content in the text or workbook. The point of these behaviors is that textbook and workbook content may not be understood completely or uniformly by all students without elaboration. This means that the instructor must enliven the content by presenting it in different forms that can highlight its most important features. The research literature suggests that oral examples, visual illustrations and practical demonstrations, if different than in the text, measurably increase the clarity of a lesson for learners who are searching for different ways to attach meaning to the lesson and relate it to their own experiences.

The research literature suggests that reviews and summaries at the end of a lesson or interspersed throughout the lesson can have much the same effect. The research also suggests that a review or summary should not be a simple parrotting of what was taught in some abbreviated

form. Instead, an effective review or summary presents content in a slightly different form than that in which it was initially presented, thereby elaborating upon and organizing it differently for efficient storage and retrieval. This provides learners the opportunity to "plug" into the content, not only at a different time but in a different manner as well. Content unlearned or misunderstood during the lesson becomes learned and clarified during summaries and reviews if the summary and review go beyond a simple repetition of the content to a repackaging of it in a manner efficient for retention and later retrieval.

Variety in the Classroom

Table 2 summarizes some of the behaviors related to instructional variety.

The research evidence. The primary and secondary sources providing strong or promising research evidence for the subdimensions of variety presented in Table 2 are as follows:

Primary

Armento, B. (1977). Teacher behaviors related to student achievement on a social science concept test. Journal of Teacher Education, 28, 46-52.

Bettencourt, E., Gillett, M., Gall, M., & Hull, R. (1983). Effects of teacher enthusiasm training on student on-task behavior and achievement. American Educational Research Journal, 20, 435-450.

Bligh, D.A. (1970). The case for a variety of teaching methods in each lesson. British Journal of Medical Education, 4, 202-209.

Brophy, J., & Evertson, C. (1976). Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.

Emmer, E. (1967). The effects of teacher use of student ideas on student verbal initiation. Unpublished doctoral dissertation, University of Michigan.

McConnell, J. (1977). Relationships between selected teacher behaviors and attitudes/achievements of algebra classes. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Ryan, F. (1973). Differentiated effects of levels of questioning on student achievement. Elementary School Journal, 41, 63-67.

Ryan, F. (1974). The effects on social studies achievement of multiple student responding to different levels of questioning. Journal of Experimental Education, 42, 71-75.

Soar, R.S. (1966). An integrative approach to classroom learning (Report for NIMH Projects No. 5-R11 MH01096 and No. 7-R11 MH02045). Philadelphia: Temple University. (ERIC Document Reproduction Service No. ED033749).

Tikunoff, W., Berliner, D., & Rist, R. (1975). An ethnographic study of the forty classrooms of the Beginning Teacher Evaluation study known sample (Tech. Rep. No. 75-10-5). San Francisco: Far West Laboratory.

Winne, P. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. Review of Educational Research, 49, 13-50.

Table 2. Effective and Ineffective Indicators for Variety

Using Variety (An effective instructor)	Poor variety (An ineffective instructor)	(Recommended behavior)
1. Uses attention-gaining devices; e.g., begins with a challenging question, visual, or example.	Begins lesson without full attention of most learners.	Begin lesson with an activity in a modality that is different from last lesson or activity; e.g., change from listening to seeing.
2. Shows enthusiasm and animation through variation in eye contact, voice and gestures; e.g., changes pitch and volume, moves about during transitions to new activity.	Speaks in monotone, devoid of external signs of emotion. Stays fixed in place entire period or rarely moves body.	Change position at regular intervals (e.g., every 10 minutes). Change speed or volume to indicate that a change in content or activity has occurred.
3. Varies mode of presentation; e.g., lectures, asks questions, then provides for independent practice (daily).	Lectures or assigns unmonitored seatwork for the entire period. Rarely alters modality through which instructional stimuli are received (e.g., seeing, listening, doing).	Preestablish an order of daily activities that rotates cycles of seeing, listening, and doing.
4. Uses a mix of rewards and reinforcers; e.g., extra credit, verbal praise, independent study (weekly, monthly).	Rarely praises or tends to use same cliches to convey praise every time.	Establish lists of rewards and expressions of verbal praise and choose among them randomly. Provide reasons for praise along with the expression of praise.
5. Incorporates student ideas or participation in some aspects of the instruction; e.g., uses indirect instruction or convergent questioning (weekly, monthly).	Assumes the role of sole authority and provider of information. Ignores student input.	Occasionally plan instruction in which student opinions are used to begin the lesson; e.g., "What would you do if..."
6. Varies types of questions (e.g., divergent, convergent) weekly and probes (e.g., to clarify, to solicit, to redirect) daily.	Always asks divergent, opinion questions (e.g., What do you think about...?) without follow-up. Or overuses convergent, fact questions.	Match questions to the behavior and complexity of the lesson objective. Vary complexity of lesson objectives in accord with the unit plan.

Secondary

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), Handbook on research on teaching (3rd ed., pp. 328-374). New York: Macmillan.

Dunkin, M.J., & Biddle, B.J. (1974). The study of teaching. New York: Holt.

Good, T., Biddle, B., & Brophy, J. (1975). Teachers make a difference. New York: Holt.

Gall, M.D. et al. (1978). Effects of questioning techniques and recitation on student learning. American Educational Research Journal, 15, 175-199.

Hanley, E.M. (1970). Review of research involving applied behavior analysis in the classroom. Review of Educational Research, 40, 597-625.

Lysakowski, R.S., & Walberg, H.J. (1981). Classroom reinforcement and learning: A quantitative synthesis. Journal of Educational Research, 75, 69-77.

O'Banion, D.R., & Whaley, D.C. (1981). Behavior contracting: Arranging contingencies of reinforcement. New York: Springer Publishing Company.

O'Leary, K., & Drabman, R. (1981). Token reinforcement programs in the classroom: A review. Psychological Bulletin, 75, 379-398.

Redfield, D.L., & Rousseau, E.W. (1981). A meta-analysis of experimental research on teacher questioning behavior. Review of Educational Research, 51, 237-245.

Rosenshine, B. (1970). Enthusiastic teaching: A research review. School Review, 78, 499-514.

Rosenshine, B. (1971). Teaching behaviors and student achievement. London: National Foundation for Educational Research in England and Wales.

Rationale for the behaviors comprising instructional variety. The following section provides an overview of the rationale for each of the subcomponents of variety based upon the above-cited research literature.

As noted in Table 2, the teaching behavior of variety involves six subdimensions:

1. Using attention-gaining devices.
2. Showing enthusiasm.
3. Varying instructional activities or teaching functions.
4. Mixing rewards and reinforcers.
5. Varying types of questions and probes.
6. Using student ideas.

The first behavior, using attention-gaining devices, is generally considered the first ingredient of a good lesson plan. Attention-gaining devices with which a lesson begins can take many forms including pictures, audio/video tapes, demonstrations, experiments; they may also have a less spectacular quality about them, such as the posing of a challenging question, presentation of a dilemma or bewildering situation, or even silence associated with a unique or interesting visual display. The research suggests that the point of beginning a lesson in this manner is to stimulate the learner in a way different than he or she has become accustomed to during the previous activity, by providing a change in the type of instructional stimuli. The point of this is to stimulate not only the receptive modalities of sight and sound but also the cognitive

processes associated with them. Without this awakening and conscious change from the mood and tempo of an earlier activity or class, the learner's attention may not be fully directed to the instructor and lesson. Attention-gaining devices, therefore, are recommended to help create natural cycles of "highs" and "lows" that make life in classrooms more interesting and less regimented to learners. Although it is commonly recognized that points of high curiosity, interest and visual impact can seldom be sustained for long periods, the research suggests that without them, there can be no feeling of anticipation which is thought to be a precondition to establishing a desirable learning set.

One of the most common recommendations from the research literature to maintain the momentum of rising and falling cycles of interest, curiosity and excitement in a classroom is to vary voice, eye contact, and gestures. When interest is waning in a class and no attention-getting device is readily available to pique the learners' attention, a change in intonation, eye contact, or gesturing is thought to be in order. Although instructors are often told not to be classroom actors, they are also told that much can be learned from the field of drama that is applicable to the classroom. For example, plays, like lessons, should have opening and closing acts and climaxes and anticlimaxes in between, to keep the audience's attention. How well an instructor plans the "script"--or lesson--has much to do with its attention-getting quality. Aside from the natural variety in topics and instructional activities a lesson may contain, variety can be enhanced by accentuating the lesson's high points and bolstering its low points with changes in voice, eye movement, and body movement. Simply put, the research literature suggests that voice inflection, eye contact, and positioning in the classroom should change often, especially during the high point (e.g., question and answer) and low point (e.g., rote recitation). Raising the voice slightly, shifting one's view from the front of the class to the back, or moving to a new location in the classroom during these times is often considered an important dimension of instructional variety.

There is little question among classroom researchers that effective teaching involves many different classroom activities. Research indicates that an instructor who does nothing but lecture for an entire period, or who only engages learners in prolonged seatwork, or who does nothing but expose learners to attention-gaining devices would have difficulty achieving successful unit outcomes, especially at higher levels of behavioral complexity (e.g., application, analysis, synthesis, and decision making). Although researchers would generally agree that some classes from time to time should emphasize a single activity, the majority of classes should offer some variety in the activities with which an instructor presents instructional stimuli. This means that some lecturing should be mixed with some questions and answers (which may include work at the board), which may be followed by guided practice in which the learners are asked to make their first attempt at responding appropriately.

These types of instructional routines, when thought out in advance and well executed, provide the variety needed for effective teaching. In this sense, the purpose of the variation in instructional activities is not so much to gain the attention of the learner but to allow the learner to cognitively process the material in a variety of different ways--for example, by listening (lectures), by seeing (textbook), and by doing (workbook). Changing the modality (hearing, seeing, doing) provides the learner with the redundancy often needed for mastery and presents the stimulus material in different contexts, emphasizing different cognitive abilities, which different learners possess to varying degrees. Therefore, a lesson plan that includes some combination of lecture, discussion, question and answer, guided practice, and independent seatwork generally is preferred to one that exclusively emphasizes only one of these instructional alternatives. Varying an instructional routine across two or more of these formats and varying the combinations chosen across days are generally believed to add the important dimension of variety to a unit plan.

One of the first things often noticed by researchers observing in classrooms is how the instructor interacts with students. Most obvious among the types of interchanges that occur and

have been reported in the research literature are those that bestow reward or punishment. Perhaps because rewards are so important in daily life, observers tend to be acutely aware of the quantity and quality of the rewards given to learners. Verbal praise is the most frequently used type of reward in the classroom. The research suggests that the reinforcing effect of verbal praise, however, depends on the variety with which it is administered and the other types of rewards with which it is associated. Most researchers agree that the phrase "That's good" or "That's correct" when repeated in the exact same voice hundreds of times across problems and learners is not reinforcing; that is, it is not perceived by the learners as a reward. The phrase soon becomes meaningless to the learner and, thereby, loses its ability to reinforce the behavior intended. The research literature points out, therefore, the importance of varying not only the type, amount and intensity of the verbal praise given but of branching out and experimenting with other types of rewards. What is rewarding to one student may not be rewarding to another. Using a learner's correct answer as a model for the next problem, having the learner retrace for the class how the correct answer was obtained, or having peers comment on the correctness of a learner's answer are some of the suggested ways that verbal praise can be varied to provide rewards that are more meaningful to the student.

Another dimension of variety which appears in the research literature pertains to questioning and probing. The art (or science) of questioning and probing is thought to be one of the most important skills of an effective instructor. It is commonly stated in the literature that the variety often conveyed to learners will to a large degree be determined by a flexible use of questions and probes. The purpose of these two related techniques is to elicit from the student a response (sometimes any response) that then can be shaped and formed into a better or more complete response. Such questions, however, are rarely considered ends in themselves but, instead, are often a precursor to subsequent learning. That is, they provide a vehicle for engaging the student in the learning process by getting the learner to act on, work through, or think about the material that has been previously presented.

Although questions are tools for engaging students in the learning process, the research suggests that to be effective, they must be administered flexibly and, often, followed up with other questions that attempt to probe into the glib, superficial or inadequate responses that are sometimes initially given to a question. Probes, therefore, are questions that follow questions, carefully crafted to deepen, enrich, and extend an earlier response.

Probes are used to elicit a clarifying response to an earlier question, solicit new information related to an earlier question, or redirect the learner into a more productive area. Each of these uses of probes is thought to offer a source of variety that can add momentum to questioning behavior, as when a solicitation for new information is followed by a request for clarification which, in turn, is followed by a redirection to a new, more productive avenue of thinking. It is recommended these cycles of probes be executed in "waves" which return again and again to force learners to act upon and reshape their responses—and most importantly, their thinking.

The research also suggests that there should be variety as well in the types of questions asked, alternating between convergent and divergent, although not necessarily to equal degrees. Convergent questions, which usually have a single right answer, are most commonly associated with the goals of direct instruction used in the teaching of facts, rules, and sequences. Divergent questions, which may have many right answers, are most commonly associated with the goals of indirect instruction used in the teaching of problem solving, inquiry, and discovery learning. The importance of these two types of questions, however, is not so much their association with direct or indirect types of instruction as it is their capacity for eliciting behavior at lower and higher levels of behavioral complexity. Convergent questioning often is recommended for eliciting behavior at the knowledge, comprehension, and application levels and divergent

questioning often is best suited for eliciting behavior at the analysis, synthesis, and evaluation levels (Bloom, Englehart, Hill, & Kranthwohl, 1956).

A final dimension of variety involves the use of student ideas. This source of variety is closely related to divergent types of questions, which have more than a single correct answer. Researchers have found that there may be no greater variety than that which occurs in learner responses to divergent questions. Every learner is an individual in his or her own right, which can make for considerable diversity in the responses obtained for divergent questions. It is the proper use of such diversity that research has suggested is linked to effective teaching. Diverse learner responses can be a problem as well as a benefit, in that unexpected or difficult-to-evaluate responses can sometimes put the instructor on the defensive. A not too uncommon response among inexperienced instructors is to invent a correct answer to end the awkwardness of not knowing how to respond to an unusual and unexpected response. According to research, this would not be a proper use of divergent questioning, since one of the purposes of divergent or "open" questions is to incorporate student ideas and participation into the lesson. The key to eliciting learner ideas and opinions is to make their ideas and opinions useful to the goals of the lesson. The research suggests that soliciting individual learner ideas and opinions seldom is considered sufficient, if it cannot be used to create greater understanding for all learners in the class.

Using this source of variety effectively means taking the contributions of individual learners and building more general concepts, patterns, and abstractions from them that are relevant to the goals of the lesson.

Task Orientation in the Classroom

Table 3 summarizes some of the behaviors related to task orientation.

The research evidence. The primary and secondary sources providing strong or promising research evidence for the subdimensions of task orientation shown in Table 3 are as follows:

Primary

Arehart, J. (1979). Student opportunity to learn related to student achievement of objectives in a probability unit. Journal of Educational Research, 72, 253-269.

Borg, W. (1979). Teacher coverage of academic content and pupil achievement. Journal of Educational Psychology, 71, 635-645.

Brophy, T., & Evertson, C. (1976). Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.

Cooley, W., & Leinhardt, G. (1980). The instructional dimensions study. Educational Evaluation and Policy Analysis, 2, 7-25.

Dunkin, M., & Doehav, S. (1980). A replication study of unique and joint contributions to variance in student achievement. Journal of Educational Psychology, 72, 394-403.

Emmer, E., Evertson, C., Sanford, J., Clements, B., & Worsham, M. (1984). Classroom management for secondary teachers. Englewood Cliffs, NJ: Prentice-Hall.

Table 3. Effective and Ineffective Indicators for Task Orientation

Being task-oriented (An effective instructor)	Poor task orientation (An ineffective instructor)	(Recommended behavior)
1. Develops unit and lesson plans that emphasize the most critical features of the curriculum or adopted text. Each unit and lesson objective can be referenced back to the curriculum guide and text.	Develops lessons mostly from personal or student interests. Breadth and depth of lesson content fail to distinguish between primary and secondary content in the curriculum guide and text.	Key each lesson to a unit plan, the curriculum guide and the text to test its relevance. Confer with other instructors concerning the most critical feature of the text and curriculum.
2. Handles administrative and clerical interruptions efficiently (e.g., visitors, announcements, dispensing of materials and supplies) by anticipating and pre-organizing some tasks and deferring others to non-instructional time.	Attends to every administrative and clerical task in detail during the time normally devoted to instruction.	Establish a maximum amount of time per hour of instruction that can be devoted to non-instructional tasks. Defer all other tasks to before or after the lesson.
3. Stops or prevents off-task behavior with a minimum of disruption; e.g., has pre-established academic and work rules to "protect" intrusions into instructional time.	Attends at length to specific off-task behavior. Lectures on the offense during instructional time.	Establish disincentives or disciplinary action for the most common off-task behaviors and post them.
4. Selects the most appropriate instructional model for the objectives being taught; e.g., primarily uses direct instruction for eliciting facts, rules and action sequences through discussion, or concepts, patterns and abstractions, through drill and practice.	Uses inefficient instructional methods for achieving lesson objectives; e.g., frequently attempts to teach facts, rules and action sequences through discussion, or concepts, patterns and abstractions through drill and practice.	Using a unit plan, curriculum guide or adopted text, divide the content to be taught into "candidates" for Type I (fact) and Type II (concept) objectives. Use direct instruction for Type I content and indirect instruction for Type II content.
5. Builds to unit outcomes with clearly definable events; e.g., weekly and monthly review, feedback and testing sessions.	Has no systematic milestones toward which to work (e.g., tests on Fridays, major review every fourth Monday) which keep the class on schedule and moving toward a clearly defined goal.	Establish a schedule in which major classroom activities begin and end with clearly visible events (e.g., minor and major tests, and review, summary and/or feedback).

Fitz-Gibbon, C., & Clark, K. (1982). Time variables in classroom research: A study of eight urban secondary school mathematics classes. British Journal of Educational Psychology, 52, 301-316.

Smith, L. (1979). Task-oriented lessons and student achievement. Journal of Educational Research, 73, 16-19.

Stallings, J., & Kaskowitz, D. (1974). Follow through classroom observation: 1972-1973 (SRI Project URU-7370). Stanford, CA: Stanford Research Institute.

Secondary

Block, J.H. (Ed.). (1971). Mastery learning. New York: Holt, Rinehart, and Winston.

Brophy, J., & Good, T. (1974). Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart, and Winston.

Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In Merlin C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 225-296). New York: Macmillan.

Duncan, M.J., & Riddle, B.J. (1974). The study of teaching. New York: Holt, Rinehart, and Winston.

Karweit, N. (1983). Time on task: A research review (Report No. 332). Baltimore, MD: Johns Hopkins University, Center for Social Organization School.

Rosenshine, B. (1971). Teaching behaviors and student achievement. London: National Foundation for Research in England and Wales.

Rosenshine, B. (1983). Teaching functions in instructional programs. The Elementary School Journal, 4, 335-351.

Rationale for the behaviors comprising task orientation. The following is an overview of the rationale for each of the subcomponents of task orientation.

As noted in Table 3, the teaching behavior of task orientation involves five subdimensions:

1. Organizing lessons in a manner that emphasizes the most critical features of the curriculum guide and text.
2. Handling administrative and clerical interruptions efficiently.
3. Stopping or preventing inattention or off-task behavior, with a minimum of class disruption.
4. Selecting the most appropriate instructional model for the objectives being taught.
5. Establishing systematic cycles of review, feedback, and testing.

Research suggests that one of the most effective ways of assuring a task orientation is to base lessons on the curriculum guide and adopted text. Although all lesson content is predetermined in the military classroom, this is not to say the effective implementation of this content in the classroom is guaranteed. Researchers have found the temptation is indeed great

among classroom instructors to "reinterpret" adopted curriculum and textbook content into topics of personal interest, topics about which an instructor may be most knowledgeable, or topics which learners appear most eager to "hear about." Although personalizing the curriculum so that it can be related to the subjective experience of learners receives support in the literature, this reformulation must emphasize the most critical features of the curriculum and text. Generally, accenting adopted curriculum topics with personal experiences, examples, and illustrations from outside the text or even outside the curriculum are considered earmarks of effective teaching if these experiences, examples, and illustrations are related in meaningful ways to the established curriculum. Developing unit plans in which personal experiences, examples, and illustrations from outside the text or curriculum are checked for their relevance to emphasizing the most critical features of the curriculum is generally considered an important step in maintaining an effective task orientation.

Another behavior considered to be an important part of an instructor's task orientation pertains to how administrative and clerical interruptions are handled. In extensive studies of more effective and less effective classroom teachers, researchers found this behavior to be among the most important in determining an instructor's task orientation for the entire class period. Researchers learned from these studies that once an instructor turned from an instructional task (e.g., lecture) to a noninstructional one (e.g., passing out materials, checking for supplies), it was increasingly difficult to return to the initial state without devoting still more time to getting learners back on task. Thus, not only was time spent on the noninstructional activity but also sometimes as much time was spent getting learners to return to their original state of attention. This has led to the recommendation to instructors to pre-organize as many clerical tasks as possible before class and establish procedures for dealing with unexpected interruptions. Taking time to collate pages of a handout, staple them and pass them out in the middle of an instructional activity, for example, often was seen as wasteful of valuable instructional time, especially when these activities could have been completed before the instruction began. Also, placing supplies and handouts on the desks of learners before class is often a recommended strategy to increase an instructor's task-oriented behavior.

A third behavior for maintaining an effective task orientation involves stopping or preventing inattention or off-task behavior, with a minimum disruption to the class. Generally, the research suggests that the significance of off-task behavior should be deemphasized at the time it occurs by the way the instructor deals with it. This does not mean that the instructor ignores or deals half-heartedly with inattention or off-task behavior, but that the instructor minimizes its consequences to the class at the time it occurs by deferring a discussion of the consequences to the student exhibiting the off-task behavior to a later, noninstructional time. In other words, the research suggests that the effective instructor deals with the offense immediately but deals with the consequences of the offense at a time and in a context chosen by the instructor. Also, the effective instructor establishes and communicates rules before the start of a course, indicating the precise procedures, disincentives and punishments to be invoked when certain offenses occur. An effective task orientation, therefore, suggests that an efficient system for handling the punishment phase of any off-task or inattentive behavior be established and communicated at the beginning of a course of instruction.

In addition to increasing the amount of class time that is devoted to instruction by pre-organizing administrative and clerical chores and establishing procedures for dealing with off-task behavior, an effective task orientation includes the efficient use of the time that is devoted to instruction. In this sense, the use of inefficient instructional strategies are often considered as wasteful of instructional time as are administrative chores and off-task behavior. Inefficient instructional strategies are thought to be especially damaging to the attainment of instructional goals. Unless considerable thought is explicitly given to the match between an instructor's objectives and the instructional activities being used to promote those objectives, the instruction may look but not actually be task-oriented.

Research on the match between an instructor's objectives and the teaching activities most appropriate for promoting those objectives has led to the development of the direct and indirect models of instruction. The functions of each are summarized in Tables 4 and 5, respectively. These models, generally, are considered to represent the means to efficiently teach two distinctly different types of behaviors. These two types of behaviors, representing the acquisition of facts, rules and action sequences and the learning of concepts, abstractions and patterns, represent the range of behavioral complexity found in almost every subject and instructional level. Researchers report, however, that it is not uncommon to observe instructors attempting to teach concepts in the context of drill and practice (direct instruction) and to teach facts in the context of an inquiry-type discussion (indirect instruction). Although both outcomes could be achieved in either context, the teaching of concepts by rehearsing facts (e.g., memorizing parts of a jet engine or reading the names on a parts list) would not be considered an efficient means of learning the concept of "thrust" or "inventory." The teaching of facts seldom allows for the generalizations and discriminations required for learning concepts; likewise, the teaching of concepts fails to emphasize the rules and sequences required for the acquisition of facts. Therefore, the importance of matching the type of learning outcome to be achieved with the instructional model which most efficiently accomplishes the outcome is considered another important behavior for establishing an effective task orientation.

A final behavior for task orientation involves establishing cycles of weekly and monthly review and testing. These cycles are built around clearly definable goals (e.g., a test at the end of the month, a review session the next week, a laboratory evaluation at the end of the unit). These are the types of "products" that research suggests should be made visible to learners and toward which instructional activities should gradually build. It has been suggested that these end products can create natural cycles of rising and falling intensity, enthusiasm and expectation, with the high point of the cycle being just before the expected event is to occur and the low point immediately afterward, marking the beginning of a new cycle. Several different cycles may occur concurrently, as when some instructional activities are preparing learners to complete a laboratory assignment or practice exercise next week, while other instructional activities are preparing learners to do well on a written test to be given at the end of the month. Thus, different cycles can be put in place for tests, assignments, and major projects, staggered in such a way as to have one cycle near its highest point when another is near its lowest. The research suggests that this can create a high level of intensity, enthusiasm, and expectation pertaining to some clearly definable goal at all times, if these cycles are made visible to students as an organizational framework for study and review.

Engagement in the Learning Process

Table 6 summarizes some of the behaviors related to a learner's engagement in the learning process.

The research evidence. The primary and secondary sources providing strong or promising research evidence for the subdimensions of student engagement behavior shown in Table 4 are as follows:

Primary

Battle, E.S. (1965). Motivational determinants of academic, task persistence. Journal of Personality and Social Psychology, 2, 209-218.

Brophy, J., & Evertson, C. (1976). Learning from teaching. Boston: Allyn and Bacon.

Crawford, J. (1983). A study of instructional processes in Title I classes: 1981-82. Journal of Research and Evaluation of the Oklahoma City Public Schools, 13 (1).

Table 4. Some Direct Instruction Functions^a

1. Reviews daily, checks previous day's work, and reteaches if necessary.
Checks homework.
Reteaches areas where there were student errors.
2. Structures and presents new content.
Provides overview.
Proceeds in small steps (if necessary), but at a rapid pace.
If necessary, gives detailed or redundant instructions and explanations.
Phases in new skills while old skills are being measured.
3. Guides student practice.
Uses a high frequency of questions and overt student practice (from teacher and materials).
Provides prompts during initial learning (when appropriate).
Allows all students a chance to respond and receive feedback.
Checks for understanding by evaluating student responses.
Continues practice until students respond quickly and accurately.
Achieves success rate of 80% or higher during initial learning.
4. Provides feedback and correctives (and recycling of instruction, if necessary).
Provides feedback to students (particularly when they are correct but hesitant).
Receives feedback from students as to whether corrections and/or reteaching is necessary.
Corrects by simplifying question, giving clues, explaining or reviewing steps, or reteaching last steps.
When necessary, reteaches using smaller steps.
5. Uses independent practice to elicit student responses that are correct and automatic.
Assigns seatwork.
Employs utilization to achieve automaticity (practice to overlearning).
Monitors students to ensure engagement during seatwork.
Achieves success rate of 95% correct or higher.
6. Provides weekly and monthly reviews.
Reteaches, if necessary.

^aAdapted from R. Rosenshine (1983).

Table 5. Some Indirect Instruction Functions

1. Provides a means of organizing content in advance.
Provides advanced organizers and conceptual frameworks, which serve as "pegs" on which to hang key points that guide and channel thinking to the most productive areas.
Allows for concept expansion to higher levels of abstraction.
2. Provides conceptual movement using inductive and deductive methods.
Focuses generalization to higher levels of abstraction by:
Inductive methods (selected events used to establish concepts or patterns), and
Deductive methods (principles or generalizations applied to specific instances).
3. Uses examples and non-examples:
To define criterial attributes and promote accurate generalizations.
To gradually expand set of examples to reflect real world.
To enrich concept with noncriterial attributes.
4. Uses questions to guide the search and discovery process; i.e., uses questions:
To raise contradictions.
To probe for deeper-level responses.
To extend the discussion.
To pass responsibility to the class.
5. Encourages students to use examples and references from their own experiences, to seek clarification and draw parallels and associations that aid understanding and retention.
Relates ideas to past learning and to students' own spheres of interests, concerns, and problems.
6. Allows students to evaluate the appropriateness of their own responses and then provides guidance as necessary.
Provides cues, questions or hints as needed to call attention to inappropriate responses.
7. Uses discussion to encourage critical thinking and help students:
To examine alternatives, judge solutions, make predictions, and discover generalizations.
To orient, provide new content, review and summarize, alter flow of information, and combine areas to promote the most productive discussion.

Table 6. Effective and Ineffective Indicators for Engaging Students in the Learning Process

Engaging students effectively in the learning process (An effective instructor)	Engaging students ineffectively in the learning process (An ineffective instructor)
1. Elicits the desired behavior immediately after the instructional stimuli; e.g., provides exercise or workbook problems with which the desired behavior can be practiced.	Fails to ask learners to attempt the desired behavior.
2. Provides opportunities for feedback in a non-evaluative atmosphere; e.g., assists learners to respond as a group or covertly the first time through. Provides needed or correct response.	Formally evaluates the initial practices; e.g., calls on learners individually to give correct answer(s) in ways that could be threatening or embarrassing.
3. Uses individualized activities (e.g., performance contracts, programmed texts, simulations, and demonstration centers) as motivational aids when needed.	Fails to match instructional methods to the learning needs of special learners; e.g., learners needing extra practice or remedial work.
4. Uses meaningful verbal praise to get and keep learners actively participating in the learning process.	Fails to provide rewards and reinforcers that are timely and meaningful to the students; e.g., never says why something is "good."
5. Monitors assignment and frequently checks progress during independent practice.	Does not monitor learner progress during seatwork evenly; e.g., spends too much time with some learners and fails to observe work of other students. Fails to monitor assignments outside of class.

Fisher, C.W., Filby, N.N., Marlave, R.R., Cahen, L.W., Dishaw, M.M., Moore, J.E., & Berliner, D.C. (1978). Teaching behaviors, academic learning time, and student achievement: Final report of Phase III-B, Beginning Teacher Evaluation Study. San Francisco, CA: Far West Educational Laboratory for Educational Research and Development.

Good, T., & Grouws, D. (1979). Teaching effects: A process-product study in fourth grade mathematics classrooms. Journal of Teacher Education, 28, 49-54.

Kouin, J. (1970). Discipline and group management in classrooms. New York: Holt, Rinehart, and Winston.

McDonald, F., & Elias, P. (1976). The effects of teaching performance on pupil learning. Beginning Teacher Evaluation Study, Phase II, 1974-1976. Princeton, NJ: Educational Testing Service.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. Educational Researcher, 8 (11), 11-16.

Wilkinson, S.S. (1980). The relationship of teacher praise and student achievement: A meta-analysis. Unpublished doctoral dissertation, University of Florida, Gainesville.

Secondary

Berliner, D. (1979). Tempus educare. In P. Peterson & H. Walberg (Eds.), Research on teaching concepts. Findings and implications. Berkeley, CA: McCutchan.

Brophy, J. (1981). Teacher praise: A functional analysis. Review of Educational Research, 51, 5-32.

Carroll, J.A. (1963). A model of school learning. Teachers College Record, 64, 723-733.

Dahloff, V. (1971). Ability grouping: Content validity and curriculum process analysis. New York: Teachers College Press.

Good, T., Biddle, B., & Brophy, J. (1975). Teachers make a difference. New York: Holt, Rinehart, and Winston.

Lysakousk, R.S., & Walberg, H.J. (1983). Cues, participation and feedback in instruction: A quantitative synthesis. American Educational Research Journal, 75, 69-77.

Rosenshine, B., & Berliner, D. (1978). Academic engaged time. British Journal of Teacher Education, 4, 3-16.

Speeth, K., & Margulies, S. (1969). Techniques for maintaining student motivation. National Society for Programmed Instruction Journal, 8, 24-27.

Rationale for the behaviors comprising engagement in the learning process. The following is a brief overview of the rationale for each of the subcomponents pertaining to engagement in the learning process.

As noted in Table 6, effectively engaging learners in the learning process involves:

1. Eliciting the desired behavior.

2. Providing opportunities for feedback in a non-evaluative atmosphere.
3. Using individualized activities and supplementary materials as motivational aids (e.g., programmed texts, simulations, technical illustrations, demonstration centers).
4. Using meaningful verbal praise.
5. Monitoring in-class assignments and checking for progress.

Engagement in the learning process begins by providing the learner stimulus material with which he/she can practice using the facts, action sequences or concepts being taught. Researchers generally agree that without practice, learning rarely if ever occurs. Researchers find that in some learning contexts it is not unusual to become accustomed to receiving lectures of an hour or longer without being provided an opportunity during the lecture to apply or to use what is being taught. At the college level, for example, where considerable independence and motivation are expected, the instructor may appropriately assume that engagement in the learning process will take place at the convenience and initiative of the learner as one works through the lecture material in the privacy of one's room or the library. However, in other learning contexts this assumption may not be valid due to the developmental level of the learner and/or the mix of other responsibilities to which the learner is expected to devote time. This possibility suggests that "lecture" and practice should go hand in hand.

Levels of development, independence and motivation generally are considered to be different for the public school learner (and also the military, vocational and technical student) than for the college student. It is generally believed that for the former, guided practice should be part of the instruction itself, since probably not all learners will know how to move from lecture to practice without the active and direct assistance of the instructor. It is often suggested, therefore, that each learner be placed in a position of grappling in trial-and-error fashion with the content being taught. It is not important that the behavior be produced at this stage in a satisfactory or correct form but rather, that the activity provided stimulates an attempt to produce the intended behavior. Such activity is intended to encourage the learner to organize a response which corresponds to the level of behavioral complexity being sought. Suggested ways in which the desired behavior can be actively elicited from the learner include oral questions, exercises from the workbook, specially prepared handout "problems," and worksheets administered by the instructor during or at the end of a lesson which take the learner through the material in a step-by-step fashion. In this manner learners are encouraged to think about, work through, or otherwise practice the material in as close a timeframe as possible to the actual instruction being provided. It is generally believed that for complex material, practice should be interspersed throughout the lesson, creating cycles of presentation and practice. Some techniques commonly recommended for providing guided practice include:

1. Asking a large number of questions.
2. Guiding students in practicing the new material, initially using prompts to lead students to the correct response and later reducing them when students are responding correctly.
3. Checking for student understanding.
4. Providing feedback.
5. Correcting errors.
6. Providing for a large number of successful repetitions.

In addition to eliciting the desired behavior through guided practice, feedback must be provided as to the appropriateness of the responses elicited, in order to get and keep learners responding. As crude, often inadequate responses are formed into slightly less crude and better responses, they, in turn, provide the basis for eventual "finely tuned" responses at a high level of accuracy. Therefore, researchers generally agree that the feedback being provided should give the learner unambiguous information as to the adequacy of any response, but in a manner that will not embarrass or humiliate the learner or lower the learner's own expectations. Since initial responses to a guided practice exercise may be crude and occasionally illogical, the potential for the student to withdraw from or diminish his/her commitment to the learning activity is at its greatest during this phase of learning. The research suggests that the instructor's role at this stage is not to pass judgment but to provide feedback in a manner that allows the learner to judge his/her own work using the guidance provided. Methods recommended for accomplishing this have included having learners respond with choral responses, showing answers on a transparency after each problem, or simply orally supplying the answer at the end of a designated time. It has been noted that it may be useful to the success of these strategies to request (and occasionally check) that learners write out their responses in order to be certain that they actually have engaged in the problem solution and not simply waited for the answer to be given.

Motivation has long been believed to be an important ingredient in the learning process. Classroom researchers have noted its importance in engaging students in the learning process. Researchers have noted, for example, that some learners do not engage fully in the learning process even when guided practice and non-evaluative feedback are provided. Some learners need "motivators" beyond these to become sufficiently excited about learning or energized to continuously respond to the practice opportunities provided. Motivational devices such as performance incentives, programmed learning texts, simulations, peer tutoring, visual documentation, and demonstration centers are generally thought useful in engaging such students in the learning process.

Foremost among these types of motivators are individualized learning materials, since they can often be used without interrupting regularly scheduled instruction. It is commonly recognized that most instructors do not have the time or flexibility to apply differing instructional methods to different types of learners. Teaching even a single group of "problem" learners needing review or remedial help, in addition to the "full" class, is not a desirable choice for most instructors. Curriculum demands prescribed by the curriculum guide or text make all such dual teaching responsibilities nearly impossible in most classrooms. However, it is generally believed that a library of individualized remedial materials can allow for different learning needs to be met for a heterogeneous class of learners. When the use of performance incentives, programmed texts, simulations, and demonstration centers is individualized, research has suggested that some students can be directed to work independently according to their own, often special, learning needs while others are working on other activities. It is believed that full engagement in the learning process by all students may require a resource library of individualized materials that can be used with special or poorly motivated learners when the need arises.

Another type of motivational device thought to be applicable to all learners is the verbal praise provided after a correct or partially correct answer. Research has suggested, however, that verbal praise in the absence of some kind of bona fide accomplishment is quickly seen by learners as an insincere response. Yet, between a glib and sometimes meaningless response such as "Correct," "OK," "Yep," and an overly emotional response such as "That's perfect," or "Great," lies a range of verbal rewards that may neither pass by barely noticed nor be taken as undeserved. These verbal rewards attempt to link a learner's response with the exact level of accomplishment attained. For example, instead of simply informing the learner that the response was "Correct," the learner might be notified that the response was "Correct because..." (e.g., the directions were followed carefully, the correct sequence of events was chosen, care was taken

to consult a reference book). Research has suggested that the motivational impact on the student, and the desire to keep engaged in the learning process, is greater when verbal praise is provided in the context of the operation successfully accomplished by the student to earn the praise. In this manner, glib phrases such as "Correct," "Good," and "OK" are never worn out, since they will be associated with some unique production by the learner; i.e., the reason the response is correct, good, or OK.

The research literature has suggested that meaningful verbal praise is especially important in the case of partially correct, or correct but hesitant responses. Here, although praise is important, it must be tempered with a sign that a better or less hesitant response is desired. To simply ignore the inadequacy or hesitation in favor of a simple "OK" is to give a sign to others that they, too, can provide a less than adequate response. In these cases, the research suggests that it is important that the praise be proportional to the adequacy of the response (e.g., "That's partly correct; now let's see if you can put it all together," or "Good try; now change the first part and you'll have it"). Examples such as these point out the often fine line between a partially correct answer and a wrong answer, which is often a matter of judgment. The recommendation from the research here is that the learners' engagement in the learning process is promoted more by choosing to see that the glass is "half-full," not "half-empty," thereby leading to praise with qualification for a partially correct answer.

Finally, the research literature has noted that rarely can a response be given that might be called "meaningful verbal punishment." Reminders to "Study harder," "Pay greater attention," "Think some more," or "Go over your mistakes" are often desirable stimulants to learning. However, phrases that ridicule, demean, or draw the class' attention to a learner's ineptitude rarely, if ever, will lead the learner to change the behavior that led to the careless or incorrect response. In these instances replacing an emotional response with constructive tips for finding the right answer is commonly recommended.

Another behavior for effectively engaging students in the learning process involves monitoring and checking during the time students are completing assignments at their seats. These two processes go hand in hand, as one is thought not to be effective without the other. Monitoring involves systematically observing all aspects of the classroom. It also means being able to perform overlapping activities; that is, actively observing learners while carrying out other activities, such as speaking and demonstrating.

During instructional activities such as guided practice and independent seatwork, it is recommended that monitoring occur within a systematic routine of checking. Rather than monitoring by simply visually scanning parts of the classroom, the instructor should circulate among students, checking responses in workbooks or exercises. The research literature is careful to point out, also, that the instructor's behavior should not be seen by learners as evaluative but rather, as a helpful response to the work being attempted. Casual glances at one paper and then another, and so on until the entire room is circulated generally are considered more constructive than long interactions with individual students. It has been recommended that interactions with individual students be limited to brief interchanges that are focused on a particular problem the student is having. This cycle can be repeated many times, in which case the instructor looks more closely the second time around at the work of those learners who were given only a glance the first time around. Monitoring and checking for progress, especially during guided and independent practice, are generally believed to represent two important tools for effectively engaging learners in the learning process.

Classes that exhibit high levels of engagement in the learning process during in-class assignments have been reported to have the following characteristics:

1. The instructor spends more time in demonstration (explanation, discussion) and guided practice.

2. The instructor makes sure students are ready to work alone, by achieving a correct response rate of 80% or higher during guided practice.
3. The seatwork activity follows directly after guided practice.
4. The seatwork exercises are directly relevant to the demonstration and guided practice activities.
5. The instructor guides the students through the first few problems.

Success Rate in the Classroom

Table 7 summarizes some of the behaviors related to establishing a moderate to high success rate.

The research evidence. The primary and secondary sources providing strong or promising research evidence for the subdimensions of establishing a moderate to high success rate are as follows:

Primary

Ames, C., & Ames, R. (1973). Teachers' attributions of responsibility for student success and failure following informational feedback: A field verification. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Bennet, N., Desforges, C., Cockburn, A., & Wilkinson, B. (1981). The quality of pupil learning experiences: Interim report. Lancaster, England: University of Lancaster, Centre for Educational Research and Development.

Brophy, J., & Evertson, C. (1976). Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.

Fisher, C., Berliner, D., Filby, N., Mariave, R., Cohen, L., & Dishaw, M. (1980). Teaching behaviors, academic learning time and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, DC: National Institute of Education.

Good, T., & Grouws, D. (1979). The Missouri mathematics effectiveness project. Journal of Educational Psychology, 71, 355-362.

Good, T., Grouws, D., & Beckerman, T. (1978). Curriculum pacing: Some empirical data in mathematics. Journal of Curriculum Studies, 10, 75-81.

Kounin, J. (1970). Discipline and group management in classrooms. New York: Holt, Rinehart, and Winston.

Zahorik, J. (1978). Teacher verbal feedback and content development. Journal of Educational Research, 63, 419-423.

Zoyonc, R., & Brickman, P. (1969). Expectancy and feedback as independent factors in task performance. Journal of Personality and Social Psychology, 11, 148-156.

Table 7. Effective and Ineffective Indicators for Moderate to High Rates of Success

Moderate to high rates of success (An effective instructor)	Poor rates of success (An ineffective instructor)	(Recommended behavior)
1. Establishes a content sequence that reflects prior learning; e.g., planning lesson sequences that consider task-relevant prior information.	Fails to sequence learning in advance to assure that all task-relevant prior knowledge has been taught before moving to next lesson.	Create top-down instructional blocks in which overall block outcomes determine order and sequencing of specific lessons.
2. Administers correctives immediately after initial response; e.g., shows model of correct answer and how to attain it, after first crude response is given.	Leaves learners to practice and learn independently immediately after presenting instructional stimuli. Waits until next day to show correct responses.	Provide for guided practice prior to independent practice and provide means of self-checking (e.g., handout with correct answers) at intervals of practice.
3. Divides instructional stimuli into small chunks; e.g., establishes "bite size" lessons that can be easily digested by learners at their current level of functioning.	Packages instruction in chunks that are too large or small; e.g., teaches too complex lesson too early in an instructional sequence.	Break lessons at "bottom" of unit plan into smaller pieces. Use original "lessons" as units or divisions within units.
4. Plans transitions to new material in easy-to-grasp steps; e.g., changes instructional stimuli according to a pre-established thematic pattern so that each new lesson is seen as an extension of previous lessons.	Abruptly changes instructional topics and perspectives from one lesson to another without themes and interconnections.	Extend unit plan hierarchy downward to more specific lessons that are tied together above with a single unit theme and outcome. Use part-whole, sequential, combinatorial, comparative, hierarchical or similar approaches for organizing content.
5. Varies the pace at which stimuli are presented and continually builds toward a climax or key event.	Maintains same pace for too long a time, leading to a monotonous and static level of intensity and expectation.	Use review, summarization feedback and testing sessions to form intervals of increasing and decreasing intensity and expectation.

Secondary

Ausubel, D.P. (1963). The psychology of meaningful verbal learning. New York: Grune and Stratton.

Cronbach, L. (1967). How can instruction be adapted to individual differences? In R. Gagne' (Ed.), Learning and individual differences. Columbus, OH: Merrill.

Lundgren, V. (1972). Frame factors and the teaching process: A contribution to curriculum theory and theory on teaching. Stockholm: Alngrist and Wiksell.

Rosenshine, B. (1983). Teaching functions in instructional programs. The Elementary School Journal, 83, 335-351.

Tobias, S. (1982). When do instructional methods make a difference? Educational Researcher, 11, 4-10.

Rationale for behaviors comprising a moderate to high success rate. The following overview provides the rationale for each of the subcomponents pertaining to success rate.

As noted in Table 7, engaging your learners in the learning process at moderate to high rates of success involves:

1. Establishing a content sequence that reflects prior learning.
2. Correcting partially correct, correct but hesitant, and incorrect answers.
3. Dividing instructional stimuli into "bite-sized" chunks at the learners' current level of functioning.
4. Changing instructional stimuli gradually, in small steps.
5. Varying the instructional tempo or pace to create momentum.

The first behavior for obtaining a moderate to high success rate occurs during unit or block planning. At that time, block outcomes are identified and a logical sequence of lessons is chosen to achieve those outcomes. Although the military instructor rarely will have the opportunity to design a block of instruction, he or she will have or should have the opportunity to influence how a block may be implemented--and thereby effect meaningful change. Researchers have indicated that some instructional blocks are planned with little or no consideration given to what must come first in a sequence of events for the block outcome to be achieved. Although some block outcomes may be achieved by lessons arranged in any order, most cannot. The result of poor lesson sequencing is often heard in expressions such as "I've run out of time," or "Had to rush to finish topic X." Depending on the circumstances, these may or may not be excuses for poor block planning but, more often than not, they could be avoided with good block planning. Researchers believe that "good block planning" means arranging lessons in a sequence that works toward long-term outcomes. This is accomplished by making each lesson "dovetail" or relate to each previous lesson. This requires that task-relevant prior knowledge be imparted immediately preceding or as close in time as possible to the lesson or lessons in which it will be needed. The savings in time results from either not having to reteach forgotten task-relevant prior learning that may be necessary for the day's lesson, or not having to reteach that day's lesson upon finding out at some later time that the task-relevant information needed for it had not been learned. Researchers note that a considerable amount of instructional time during a course of instruction is sometimes consumed by backtracking to remedy deficiencies in facts, skills or

concepts that could have been avoided with better planning at the unit level. Effective block planning means making each lesson work to accomplish the goals of subsequent lessons by arranging a sequence of instruction that builds to block outcomes gradually, in a logical and systematic order. The recommendation most commonly made here is to make each new lesson a logical extension of the previous lesson in order to secure moderate to high success rates on performance exercises during guided and independent practice.

Another behavior thought essential for securing moderate to high success rates on tests, homework, and problem sets is the provision of timely feedback. During guided practice, in which the desired behavior is being elicited for the first time, correctives should be given immediately after the learner's initial response. The time between practice and feedback has long been considered in the research literature as one of the most important elements of learning. It is generally agreed that the longer that feedback is delayed, the less likely it is to influence the learner's performance on subsequent attempts to produce the behavior. The reasons provided for this are that the learner must hold a mental image of the first crude response in order for the feedback to be effective. Unfortunately, mental images do not last long after an initial response is made; so, the effectiveness of the feedback deteriorates rapidly with any delay. In order for the learner to link feedback with his/her image of the response, the corrective must follow immediately. Recommendations for accomplishing this often include calling out the right answer after each practice item has been completed, displaying on a transparency a model for attaining a right answer, or having learners check the responses of other learners according to some standard provided by the instructor or the text. The research notes also that feedback using these and related procedures should be administered in a non-evaluative atmosphere. Both immediacy and a non-evaluative atmosphere are recommended in order for the feedback to have its greatest effect on revising and eventually fine-tuning the response.

In order to achieve moderate to high rates of success, the research suggests that feedback should be provided for each and every item given during guided practice. Missing the opportunity to correct a wrong, partially wrong, or even correct but hesitant response (which may be correct for the wrong reason) during this critical time can affect success rates during later independent practice (in the form of sustained work at desk, in workbook exercises or on homework). Research suggests that guided practice should continue until a success rate of approximately 60%-80% is achieved. Generally, that will be possible only if the correct responses and the reasons for them are provided after every trial response. This moderate to high success rate is then believed to pave the way for still higher success rates at the end of independent practice, at which time about 90% correct answers should be expected. The research suggests, however, this higher rate of success can rarely be achieved without first establishing a 60% to 80% success rate during guided practice. Finally, some research suggests that when a 60% to 80% success rate is achieved during guided practice, feedback during independent practice may be delayed with no ill effects on success rate, as when correct answers are given at the end of the period or homework is graded the next day.

The next behavior often considered important for establishing moderate to high success rates involves tailoring block and lesson plans to fit the learners. To be sure, blocks need to be chosen to reflect some "whole" that is comprehensible to learners. But, the research also points out that blocks must be sized to cover a length of time in which learners still remember the specifics taught at the beginning of the block at the time when broader concepts and abstractions are taught at the end of the block. Long blocks that ramble through, for example, many physical laws, mathematical procedures, logistical operations, rules, and regulations--no matter how related the laws, procedures, operations, rules and regulations may be--may prevent some types of learners from connecting facts, rules, and sequences with the concepts, abstractions and patterns to which they belong. Research suggests that the pace and size of some blocks may have to be adjusted for some types of learners and some types of subject matters.

Inappropriate sizing at the block level has been reported to create serious problems at the lesson level. For instance, when a block topic is too broad, the typical instructor tends to make lessons more abstract in order to cover the block content in the specified period of time. As a result considerable difficulties occur at the lesson level in establishing moderate to high success rates. As lesson content becomes less specific, adequate periods of guided and independent practice are often slighted. There is a rush to cover the content, with perhaps some compromises made on how well it is covered. It has also been found that when this happens, behaviors at higher levels of complexity may be taught without adequate attention to the task-relevant prior facts, rules, and sequences that may be necessary for attaining these higher-level outcomes. It has also been noted that when instruction is continually planned around the indirect instruction model (see Table 5), it may be a sign that the content area is too broad at the block level and that some retrenchment may be necessary. Although the determination of block content is not the responsibility of the instructor, practical decisions about the implementation of this content, if properly recorded, can from time to time make valuable adjustments to a block, which can be formally changed at a later time. The research suggests that every effective instructional environment should encourage this type of individual initiative and flexibility.

Finally, it has been noted that the size of the chunk chosen for lesson content must be at the learners' current level of functioning, if moderate to high success rates are to be achieved. When guided practice sessions or question and answer sessions do not provide 60% to 80% correct responses, lesson content may be both too broad and too complex. This is commonly taken as a sign that chunks of content are not "bite sized" and that the lesson may have to be divided into several smaller segments before it can be digested by learners.

A fourth behavior for establishing moderate to high success rates involves transitions within and between chunks of lesson content. Although planning lesson content in bite-sized chunks aids learners in making transitions between old and new content, the research on planning suggests this alone will not be sufficient to assure moderate to high success rates. In addition, the instructor needs to present block and lesson content in ways that establish some overarching themes which interconnect different parts of a block or an extended lesson. Although not mutually exclusive, such methods include the part-whole, sequential, combinatorial, comparative or hierarchical approaches to organizing content that are commonly described in teaching methods texts (Gage & Berliner, 1984). Each of these approaches represents a means for organizing lessons, and content within lessons, that emphasizes relationships among various aspects of a topic. Some of these approaches link content together procedurally, as in the sequential approach, while others link content by emphasizing themes and threads running through the content itself, as in the part-whole approach. The research suggests that the effect upon the learner is to make the transitions between parts of a topic more comprehensible; that is, it helps the learner see that what is being taught now is actually a part of what has gone before and what will follow. This aids in establishing moderate to high rates of success on performance assessments inasmuch as the learner is able to build an understanding of the topic gradually in measured steps, as opposed to trying to put all the bits and pieces together at review and test times. Easy and hard lessons may be inevitable, but the research suggests that transitions between them will be more comprehensible when the common structure of which they are a part is made known to learners.

Another behavior for establishing moderate to high success rates has been mentioned previously in relation to several other teaching behaviors. It deserves, however, special mention again with respect to the concept of momentum. This teaching effectiveness indicator adds a subtle but important feature to establishing cycles of weekly and monthly review and testing, as discussed earlier with regard to task orientation. To illustrate the concept of momentum, some writers have compared the practice of teaching with that of conducting a symphony orchestra. The thread which binds these two activities together, it is pointed out, is not

content but rather the simple fact that both the maestro and instructor must have a pattern or theme that is understood by their respective audiences even if it has never been heard before. How learners can be made to understand in the course of a 60-minute lesson something they knew nothing about before, and how the conductor can make an abstract and perhaps-never-heard-before concerto recognizable to the audience, is a topic that has gained the attention of a number of writers. Their tentative conclusions suggest that the explanation lies in the way the individual notes or individual pieces of the lesson are put together. How the notes are played against one another is reflected in the tune's melody and how the individual pieces of content are related to one another are reflected in a lesson's momentum. The literature on effective teaching generally agrees that one of the most important ingredients for ensuring that a lesson will be listened to and understood is to write one that has momentum. Momentum is the varying pace at which stimuli are presented in the process of reaching a climax or key event. In teaching, it is suggested that this be accomplished first by establishing cycles of weekly and monthly review, feedback, and testing, and then, by gradually increasing the instructional pace and intensity as the time for the major event draws near. Playing the same instructional "note"--or keeping the same monotonous pace too long--may be as boring as listening to a drab and lifeless musical score. Consequently, rising and falling actions are established to set the instruction in motion toward some discernible event. The event can be established by scheduling a cycle of review, feedback and testing and by gradually increasing the pace or tempo within the cycle until the end of the cycle is reached (e.g., a test or laboratory evaluation), at which time the cycle can begin again.

Some techniques commonly recommended for securing moderate to high success rates include:

1. Break down the instruction into smaller steps. Give the students instruction and practice on each step before proceeding to the next step.
2. Provide the students with explicit demonstrations of skills, whenever possible.
3. Intersperse demonstrations with questions in order to maintain students' attention and to check for student understanding.
4. Provide students with instructor-monitored practice prior to seatwork activity so that the instructor can correct errors before they become part of the students' repertoire.
5. With especially confusing material, provide pre-corrections by advising the students about particularly confusing areas.
6. Provide sufficient independent practice, in terms of both length and number of exercises, to enable students to master skills to the point of overlearning (with additional exercises for the slower students).
7. Reteach material when necessary.

III. INSTRUCTOR EVALUATION PROCEDURES IN THE PUBLIC SCHOOL SECTOR

This section briefly reviews some of the evaluation and feedback systems currently in use by school systems and State Departments of Education to evaluate and provide feedback to public school teachers. The teaching effectiveness indicators used in one of these systems are compared with those from the research literature cited in the previous section.

Throughout much of the past decade, the U.S. has engaged in a major movement of educational reform. The primary impetus for this movement has been the release of several reports critical of the American school system, such as the often-cited report A Nation at Risk, released in 1983

by the National Commission on Excellence in Education. The seeds of dissatisfaction which in the 1960's and 70's led to competency testing requirements for students have since burgeoned into a mandate for comprehensive educational reforms touching on a number of new initiatives pertaining to the evaluation and retention of both students and their classroom teachers. For example, 27 states to date have passed major educational reform packages pertaining to required standards of performance for students and teachers (Hardy, 1986).

Two related developments have proceeded from these new initiatives. One has been the spreading use of teacher competency testing. The second has been an increase in the creation, refinement, and use of teacher appraisal instruments, which include direct observation of on-the-job classroom performance. Each of these developments is briefly reviewed below.

Competency Testing

Between 1977 and 1985, the number of states requiring some form of teacher competency testing increased from 3 states to 42. Testing programs vary widely from state to state, however, differing in their target populations, test content, and use of results.

Target population. At least a dozen states test students prior to admission to teacher education programs. The most common (and nearly universal) target population for competency testing is the set of prospective teachers at or near completion of coursework who are attempting to obtain certification. Only three states (Texas, Arkansas, and Georgia) reported the use of competency testing with teachers who are already certified (Lehmann & Phillips, 1987).

Test content. The most frequently employed competency tests measure knowledge of basic skills such as reading and mathematics. Hardy (1986) reported that 18 states require subject area specialization tests (e.g., prospective secondary science teachers are tested on secondary science content). Additionally, general knowledge of a professional nature (such as teaching methodology) is tested by 19 states. Hardy pointed out that states often employ a combination of basic skills, subject area specialization, and/or professional knowledge testing. Two of the most popular competency tests currently in use are the National Teacher Examinations (Core Battery and Specialty Area Exams), and the Pre-Professional Skills Test.

Test administration and use of results. Competency tests are most commonly administered either by universities or by agents of State Departments of Education, and are primarily used for the purpose of initial certification, rather than for salary or promotion decisions. Florida and Tennessee have utilized competency data in master teacher or career ladder programs, but this is not the norm. Utah is currently allowing National Teacher Examination scores as one line of optional evidence teachers may use in seeking promotion. Competency data have also been used as a means of assessing the quality of teacher education programs in at least eight states.

Trends. The use of competency testing as a requirement for teacher certification continues to increase. Haney and Reidy (1987) cited three major initiatives at the national level which have appeared in only the last 3 years and which will likely add to the momentum of the teacher examination movement:

-- As of 1985, the National Council for Accreditation of Teacher Education (NCATE) has required teacher training institutions to assess (a) the basic skills of candidates for entry into teacher training, and (b) exiting student competencies, as a requirement for eligibility for evaluation for accreditation.

-- The roughly 100 universities involved in research and teacher training who are participating within the consortium known as the Holmes Group in an effort to reform

teacher education, in 1986 called for the creation of improved entry standards for the teaching profession. Although the Holmes Group is "seeking ways to broaden the nature and scope of the assessment of critical characteristics and knowledge in the preparation of a teacher" (Cole, 1987, p. 25), acknowledgment is made of the need for demonstrated basic skills, subject area and professional knowledge.

-- The Stanford Teacher Assessment Project's development of innovative measurement devices is in answer to a need identified by the Carnegie Forum on Education for a "National Board of Teaching Standards." The Project has been funded for the purpose of providing prototypical teaching assessment devices for the National Board.

As the use of competency exams for prospective teachers spreads, the trend appears to be toward the use of more sophisticated measures employing non-paper-and-pencil sources of data that can ensure that those entering the field of teaching can actually demonstrate their knowledge and skills in the classroom.

Performance Appraisal

A second educational reform initiative has led to the increased assessment of teachers already in the field. Prior to the past 10 years, appraisal policies for practicing teachers were largely under the control of the local school districts in most states. This continues to be the case in many states. However, there is a growing trend toward state-level involvement. States such as Arkansas, Florida, Kentucky, Oklahoma, Tennessee, and West Virginia have developed guidelines which local districts must follow in developing teacher evaluation instruments and/or procedures. Moreover, some states have mandated the use of statewide instruments and procedures for evaluation of all teachers. Included among the states which have committed man-hours and money to the development of comprehensive, state-of-the-art evaluation systems for in-service teachers are Georgia, Texas, Virginia, North Carolina, Mississippi, and Delaware.

In their recent survey of the teacher evaluation programs of the nation's 100 largest school districts, Ellett and Garland (1986) found a significant qualitative discrepancy between local district systems and the comprehensive statewide systems. Accordingly, the remainder of this section will report first the findings of Ellett and Garland regarding the practices and policies of the large local districts surveyed. This will be followed by a more detailed description of Kentucky's guidelines for teacher performance evaluation (as an example of a state system which provides such guidelines), and then by an analysis of the statewide procedures and instruments in use in Mississippi and Texas.

Performance appraisals by large school districts. Ellett and Garland requested information from the United States' 100 largest school districts regarding the purposes, policies, and practices of their teacher evaluation systems. Their reported data represent a return rate of 80%, with no information supplied regarding possible systematic differences between responding and non-responding districts.

When asked to rank order the following four possible purposes of their teacher evaluation systems, respondents' mean rankings reported "professional development for teachers" to be of highest importance, followed in descending order by "accountability," "personnel decisions," and "instructional leadership for administrators." Reports on the actual utilization of evaluation data reveal somewhat of a discrepancy, however. The most commonly reported uses (selected by over 90% of the respondents) were (a) "development of remediation plans for teachers with identified deficiencies," and (b) "teacher dismissal." "Promotion/compensation" and "merit pay" were indicated as uses for appraisal data by 50% and 23.8% of the respondents, respectively. In

actual practice, teacher evaluation data appear to be used for both professional development and evaluation, although almost twice as much for the former as for the latter purpose.

A formal written policy addressing teacher evaluation was reported to be in effect in 97.5% of the responding districts. Practical issues, such as who is evaluated by whom and when, are included in at least 90% of the policies; other issues are less frequently addressed. "Standards for acceptable teaching" and "orientation of evaluators to the evaluation instrument" are stipulated by 70.7% and 61.3% for these policies, respectively. Although 91.3% reported that evaluation procedures are clearly written in detail, only a fourth of the districts actually require potential evaluators to demonstrate mastery of the use of the evaluation instrument. Finally, less than half the districts reported policy provisions for "systematic review and revision of teacher evaluation policies."

As noted in Table 8, the most commonly cited method of evaluating teachers was "direct, systematic observation of teachers." Of those districts employing this method, over 70% utilized both a standardized observation form and timelines for completing the observation. Over 88% of these districts reported the use of post-observation conferencing between the evaluator and the teacher. Rarely was the procedure for combining the data from more than one observation on an individual spelled out, with only 26.6% reporting the existence of written rules for such a process.

Table 8. Methods of Evaluation Reported by the Nation's 100 Largest School Districts^a

Methods	Percent
Direct, Systematic Observation of Teaching	98.8
Informal Observation of Teachers	85.0
Peer Ratings of Teacher's Performance	11.3
Student Ratings of Teacher's Performance	3.8
Student Achievement Data	17.5
Paper-and-Pencil Examinations	10.0
Teacher Self-Evaluation	31.3
Other	8.8

^aEllett and Garland (1986).

In addition to the quantitative analysis of the survey data received, Ellett and Garland also conducted a qualitative analysis of the supporting documentation received from 30 districts. Twenty-five districts furnished evaluation instruments and/or policy descriptions. The structure of the instruments showed from 10 to 65 performance criteria being listed under from 2 to 25 broad categories of teacher attitudes. The performance criteria were explained in detail or clarified by behavioral indicators in less than half of the instruments received. Unfortunately, although Ellett and Chandler gave examples of the broad categories employed ("Preparation and Planning," "Classroom Organization and Management," and "Professional Growth and Attitudes"), they gave no detailed information regarding the specific criteria or behavioral indicators found on these instruments. They did note, however, that few of the instruments were designed on the basis of teacher effectiveness research, and that most included assessments of high-inference variables. They stated:

...many school districts' teacher evaluation instruments depend totally upon personal judgments of evaluators who receive little or no systematic training ... relatively few criteria were at a level of behavioral specificity necessary to train observers to make reliable assessments of teacher performance ... it is apparent that translation of state-of-the-art practices in teacher evaluation from large-scale assessment applications to local school district uses seems sorely lacking. (p. 26)

Although Ellett and Garland's analysis of the state of the art of teacher evaluation at the school district level may be overly critical, as the following data will suggest, it probably does accurately describe conditions within at least some school districts.

Performance appraisals by State Departments of Education. The development of statewide guidelines or mandates for teacher evaluation in the U.S. has been intimately tied to the creation, by many states, of teacher incentive programs, such as career ladder programs or bonuses for outstanding performance. Of the 12 states previously identified as having statewide guidelines and/or instruments, all but Oklahoma have state incentive programs which are either fully implemented, in the pilot-testing stage, or otherwise in the process of development. A recent count (Cornett, 1986) placed the number of states involved to some extent in incentive programs at 29. The evolution of incentive systems has had significant implications for teacher evaluation. In an update on trends in teacher incentive programs, Cornett reported that states with incentive programs are departing from traditional, relatively subjective, global ratings by principals. Emerging program characteristics cited include research-based instrumentation, the inclusion of multiple sources of evaluation data, and greater emphasis on evaluator training. Three specific state evaluation systems are considered below:

Kentucky. In March 1985, the Kentucky Board of Education issued its Guidelines for Teacher/Administrator Performance Evaluation in Kentucky, as required by Kentucky Senate Bill 364 enacted in 1984. The Guidelines, as stipulated by legislation, require (a) that evaluation procedures be developed by certified teachers and administrators, (b) that evaluation include an evaluator/evaluatee conference, (c) that evaluators be trained in evaluation techniques both general and specific to the local system, and (d) that the procedures developed include plans to aid each evaluatee in increasing his or her effectiveness. All evaluators are trained and tested in 4-day sessions.

The Guidelines require that the following performance criteria be included in the evaluation instrument. (Others may be added at the local school district's discretion.)

1. Performs professional responsibilities and duties as outlined in the job description, with regular attendance and punctuality.
2. Uses instructional strategies and processes effectively.
3. Demonstrates effective interpersonal and communication skills with peers, subordinates, students and parents.
4. Plans and evaluates instructional activities.

Also required are behavioral indicators which define the criteria. Each indicator is to be written such that acceptable quality of performance is specified. The Guidelines include examples of matched criteria and indicators, as supplied by Dr. Jim Sweeney, Iowa State University. These may be found in Appendix A (specimen 1). Also included in the Guidelines are the teacher evaluation instruments of Danville and Covington, Kentucky. On these forms, each indicator consists of a range of observed behaviors which are assigned ratings. The Danville form lists performances for each indicator which may be rated unsatisfactory, improvement needed, competent, very good, or outstanding. The Covington instrument assigns 3, 2, 1, or 0 points to each level of performance. The summary forms for recording evaluation data from an observation in Danville and Covington are also provided in Appendix A (specimens 2 and 3).

Mississippi. Use of the Mississippi Teacher Assessment Instruments was fully implemented in the 1986-1987 school year. Results of beginning teachers' performance evaluations are to be used in state approval decisions concerning teacher training programs. Additionally, results of

all teachers' performance evaluations will determine each teacher's eligibility for a \$1,000 incentive bonus to be awarded in 1987-1988. Documentation from the Bureau of School Improvement indicates the use of three evaluators: one external, the building principal, and a peer teacher. This policy is targeted at beginning teachers. Cornett (1986) reported, however, that for incentive bonus purposes, only trained administrators will act as evaluators. Provisions have been made for remedial training of teachers who do not qualify for the merit pay.

The three instruments which comprise the Mississippi Teacher Assessment Instruments measure a total of 14 "competencies" which are defined by 42 indicators. Appendix A (specimen 4) provides a list of these competencies and indicators. Indicators are rated from 1 to 5, with detailed descriptors provided to clarify the meaning of each scale point. An example of such descriptor information is included with specimen 4 in Appendix A. Competencies I to IV are assessed with the Teaching Plans and Materials (TPM) instrument. Information to be used by the evaluator(s) in completing the TPM comes from a portfolio and a questionnaire, both completed by the evaluatee. In contrast, the evaluator is to use on-the-job observation in the completion of the Position Skills and the Interpersonal Skills instruments.

Texas. Texas legislation in 1984 established a career ladder incentive system for teachers and required the development of a statewide evaluation plan and criteria. In 1984-1985 and in 1985-1986, the career ladder employed locally developed evaluation instruments. The statewide plan was fully implemented in 1986-1987, after being pilot-tested in six districts the preceding year.

The Texas Teacher Appraisal System (TTAS) policies are comprehensive. Title 19, Part II of the Texas Administrative Code and Statutory Citations, addresses in detail issues of appraiser qualifications, teacher orientation, observation and conferencing processes, teacher response and appeals, professional growth plans, teacher self-appraisal, summative appraisal, and scoring procedures. Notable points include requirements for a minimum of six program-orientation hours for teachers, required pre- and post-conferences, and written quantitative procedures for combining the results of multiple observations and for assigning qualitative rankings to summative scores.

The Teacher Orientation Manual describes the process of literature review and teacher input which were part of the development of the TTAS criteria. The manual's appendix lists the relevant research references. The TTAS was intended to have a narrow focus on criteria which could be validly and reliably measured, and which were identified by research as being significant.

The four classroom teaching domains thus identified were the following:

1. Instructional Strategies
2. Classroom Management and Organization
3. Presentation of Subject Matter
4. Learning Environment

An additional domain, Growth and Responsibilities, is included for which full credit is given in the absence of any documentation of unacceptable practices. Local districts may develop additional domains and criteria, but results of these will not impact career ladder decisions. Appendix A (specimen 5) includes the observation record which allows indicators to be scored directly, as well as the appraisal record on which are documented the domain scores from several observations and which shows the calculation of the overall summary performance score.

Finally, it can be noted that when the indicators of teaching effectiveness from the research literature reviewed in Section II of this report are compared with those indicators listed on the Texas Teacher Appraisal Instrument, a 79% correspondence occurs. That is, 79% or about 22 out of the 28 indicators of teaching effectiveness identified in Tables 1, 2, 3, 6, and 7 of Section II have related or corresponding items on the Texas Teacher Appraisal Instrument. Conversely, only approximately 56% of the teaching effectiveness indicators on the Texas instrument have corresponding or related entries in Tables 1, 2, 3, 4, and 7, indicating the inclusion on this instrument of both research-based and a large number of professionally based behaviors typical of appraisal devices in the public school sector.

IV. DESCRIPTION OF INSTRUCTOR EVALUATION FORMS IN THE AIR FORCE, ARMY, AND NAVY AND HOW THEY ARE USED IN THEIR INSTRUCTIONAL ENVIRONMENTS

This section reviews the current instructor evaluation procedures used by the Air Training Command (ATCR 52-8) and compares the teaching effectiveness indicators on ATC Form 281 with those derived from the research literature. This section also briefly summarizes the instructor evaluation and feedback procedures currently in use by the Army and Navy, with emphasis on commonalities among the three services.

Instructor Evaluation in the Air Force

The evaluation of instructors in technical training centers in the Air Force is comprised of five distinct types of data and related procedures:

1. ATC Form 281, Instructor Evaluation
2. ATC Form 736, Student Critique
3. Training Evaluation Report
4. Student Exercises and Tests
5. Informal Peer Evaluation

ATC Form 281: Instructor Evaluation. This form contains 35 items, divided among 12 general factors. The 12 general factors measured are:

Preparation (4 items)
Content/Structure of Presentation (3 items)
Presentation Skills (3 items)
Communication Skills (3 items)
Demonstration Skills (2 items)
Use of Training Aids/Materials (2 items)
Question/Answer Techniques (4 items)
Management (5 items)
Supervising Group Activity Skills (2 items)
Measurement (2 items)
Personal Qualities (3 items)
Other (2 items, with an option to add others)

Each item on Form 281 is rated on a 4-point scale as either Outstanding, Satisfactory, Needs Improvement, or Not Applicable. Space is provided in the right-hand margin of each item for

comments. Raters are advised on the top of the form that an "Outstanding" or "Needs Improvement" rating for any item requires justification in the comment space provided and that, additionally, a "Needs Improvement" rating must be followed up with a subsequent evaluation within 30 days. In addition to the 35 item scores, a single overall performance rating is given, using the scale points of "Outstanding" (when more than 75% of the items are rated "Outstanding" and no item is rated "Needs Improvement"); Satisfactory (when the majority of items are rated either "Satisfactory" or "Outstanding" and less than 25% are rated "Needs Improvement"); "Needs Improvement" (when more than 50% of the items are rated "Needs Improvement"). Space is also provided for recommendations for improvement and/or general comments and for comments pertaining to a follow-up evaluation, should any item be rated as needing improvement.

The instructor's designated supervisor completes Form 281 after observing the instructor presenting a typical lesson for a minimum of 50 minutes in the classroom, the laboratory or the field. New instructors must be evaluated every 30 days for 3 months with Form 281 and, then, once each quarter until the instructor attains the designation of master instructor, at which time evaluation is yearly. The results of Form 281 are conveyed to the instructor in a conference setting shortly after the evaluation, at which time it is customary for the supervisor to review the instructor's strengths and weaknesses, with particular reference to items checked "Outstanding" or "Needs Improvement" and the comments which have been made to justify these ratings. Visits to ATC classrooms and post-instruction interviews with instructors at Lackland AFB, Keesler AFB, Lowry AFB, and Sheppard AFB by the contractor during the months of June and July, 1987, suggest that:²

- Overall, instructors tend to value the feedback provided them by Form 281.
- Some feel a need for more specific feedback, particularly with respect to items on which they have been rated "Satisfactory," but for which they wish to attain a rating of "Outstanding."
- Little variation in ratings occurs either within or between individuals; most instructors receive a rating of "Satisfactory" on most items.
- Generally, instructors believe Form 281 is more appropriate for formal classroom instruction than for laboratory work, performance classes, or field exercises, which tend to require instructional behaviors more specific to these instructional contexts.
- Most instructors would like to receive more written support, recommendations and justifications in the comments section of Form 281, and most instructors would like to see more space allotted to comments on the form.

ATC Form 281 appears in Appendix C.

ATC Form 736: Student Critique. This is a single-page, open-ended form for students who wish to voluntarily evaluate the training course and/or instructor. Students are instructed to use this form to evaluate their training, student group or base support facilities and services, but to use a separate Form 736 for each of the areas being evaluated. Additionally, Form 736 includes space for the course title and number, period of training, focus of the evaluation (training, school squadron, or base support facilities), and type of critique (individual, group or oral-type). The student may request to receive the results of his critique, in which case a check is placed in an appropriate box. Form 736 also includes a space for the instructor to identify any follow-up action taken and the method of contacting the student when the results of

²See Appendix B for the specific ATC classrooms observed by the contractor at these bases.

the critique are requested by the student. Form 736 may be submitted individually or a single Form 736 may be submitted by an entire class. In the case of the latter (required by some schools and instructors), a predesignated student solicits oral comments from the class during a specially allotted time in which the instructor is absent. Generally, comments made by individual students with which 50% or more of the class agrees are recorded anonymously on Form 736. Visits to ATC classrooms and post-instructional interviews at Lackland AFB, Keesler AFB, Lowry AFB, and Sheppard AFB suggest that:

1. Form 736 tends to be most valued by instructors for, and is most frequently used by students for, identifying specific, well-defined weaknesses of a course or instructor. Form 736 tends to be less valued by instructors for, and is less frequently used by students for, documenting good or exceptional teaching performance, in which case student comments tend to be congratulatory and instructionally vague.
2. Instructors tend to treat student requests for feedback on Form 736 seriously, frequently getting in touch with individual students to discuss complaints, when possible.
3. Overall, Form 736 is most used when identifiable problems persist in the classroom and students believe there is a willingness on the part of the instructor to accept feedback. In classrooms where this atmosphere is not communicated, Form 736 tends not to be used by individual students.
4. The number of individually submitted 736's varies considerably from instructor to instructor, with some senior instructors reporting never having received one, and other instructors reporting having received some each time their course was taught. Generally, the number of individually submitted Form 736's appears not to be large.

Form 736 appears in Appendix C.

Training Evaluation Report. The training evaluation report is a post-instruction questionnaire which queries graduates on the job about the adequacy of their course instruction. A small set of graduates of each training school are randomly sent a questionnaire approximately 6 months after training, while performing the duties related to their occupational specialty. The questionnaire provides students with an opportunity to comment on the adequacy of the training received from individual courses as well as instructors. Most instructors and some supervisors, however, cannot remember receiving results from the questionnaire and, hence, it generally is not considered an evaluative or diagnostic tool for instructor feedback. From the instructors' perspective, most of this information remains at the command level for administrative purposes.

Student Exercises and Tests. These include both informal exercises and formal tests of student knowledge given in the context of instruction. Student exercises include practice exercises, routinely a part of most lessons, whereas tests include both the results of performance evaluations that are graded within a course and formal tests at the completion of the course. Typically, instructors report using the results of these exercises and tests for self-evaluation and, when necessary, to revise instructional procedures, although on an informal basis. However, instructors report that student data in the form of responses on exercises and grades on tests, although reported to be useful as general benchmarks of teaching performance, do not provide specific diagnostic information as to what teaching behaviors might be related to a good, mediocre, or poor class performance. Therefore, the informal use of these data for instructor evaluation and their lack of diagnostic specificity with regard to specific instructor behaviors appear to limit their usefulness for feedback enhancement.

Informal Peer Evaluation. A final source of data for the evaluation of training is provided by peer evaluations, in which instructors voluntarily observe each other. The frequency of peer evaluation varies from school to school, with most schools including some peer evaluation and a few schools scheduling frequent evaluations as a standard practice. Peer evaluations most commonly use Form 281, which either is completed by the peer observer, just as it would be completed by the supervisor, or is used as a general guide for what to observe. Feedback is informal, many times taking the form of brief chats between instructors immediately after the observation or queries in the breakroom as to "What did you see?" Visits to a sample of ATC classrooms at Lackland AFB, Keesler AFB, Lowry AFB, and Sheppard AFB suggest that:

1. Peer evaluation is valued by instructors as one of the most useful sources of feedback they receive.
2. Peer evaluation often occurs too infrequently, due to time and scheduling commitments.
3. Peer evaluations tend to be the most helpful when the peer evaluator looks for instructional effectiveness in ways that supplement or enhance items on Form 281; i.e., makes extensive comments in the space provided.
4. Although Form 281 often serves as a starting reference for a peer evaluator, observers often add to it using evaluative criteria of intuition, professional judgment, and personal experience. Often such additions appear in the comments portion of Form 281.

Comparison of Items on ATC Form 281 with Indicators of Teaching Effectiveness from the Research Literature

In this section the items on Form 281 are compared with the indicators of teaching effectiveness reviewed in Section II of this paper. In that section five key behaviors (Clarity, Variety, Task Orientation, Engagement in the Learning Process, and Moderate to High Rates of Success) were identified from the research literature. These five key behaviors, in turn, yielded 28 different subdimensions of teaching effectiveness that were identified in Tables 1, 2, 3, 6, and 7. In Table 9, the 35 items on Form 281 are "mapped onto" these 28 behaviors in an effort to determine the degree to which current Air Force instructor evaluation criteria reflect the criteria for teaching effectiveness commonly cited in the research literature.

Of the 35 behaviors on ATC Form 281, 28 items or 80% have some degree of correspondence with the indicators from the teaching effectiveness literature. Those ATC Form 281 items for which no counterpart could be found were:

- A2 Classroom neat and orderly. Seating arrangement appropriate. Items identified needing repair, such as burned out lights.
- D1 Correct enunciation and grammar used.
- H4 Instructor properly used the MIR instructor.
- K2 Instructor's professionalism set the proper example for bearing, behavior, and dress.
- K3 Positive rapport with students established.
- L1 Students in compliance with appropriate dress and appearance.
- L2 Importance of safety emphasized and compliance with safety standards ensured.

Table 9. Comparison of Key Indicators of Teaching Effectiveness from the Research Literature with Items on ATC Form 281

Key behavior	Teaching effectiveness indicators from research literature	Related items on Form 281
Clarity	<ol style="list-style-type: none"> 1. Informs learners of skills or understandings expected at end of lesson. 2. Provides learners with an advanced organizer with which to place lesson content in perspective. 3. Checks for task-relevant prior learning at beginning of lesson, and reteaches when necessary. 4. Gives directives slowly and distinctly. Checks for understanding along the way. 5. Knows learners' ability levels and teaches at or slightly above their current level of functioning. 6. Uses examples, illustrations or demonstrations to explain and clarify content in text and workbooks. 7. Ends lesson with review or summary. 	<p>B. 1 Objectives clearly stated and sequences for lesson briefly outlined.</p> <p>B. 1 Clear instructions provided to the group.</p> <p>E. 1 Skills properly introduced and demonstrated.</p> <p>B. 3 Main points reviewed in conclusion (or internal summary given if lesson not ended).</p>

Table 9 (Continued)

Key behavior	Teaching effectiveness indicators from research literature	Related items on Form 281
1. Uses attention-gaining devices.		
2. Shows enthusiasm and animation through variation in eye contact, voice, and gestures.	C. 1 Eye contact made with students. C. 2 Movement and gestures natural and appropriate, not forced or distracting. C. 3 Instructor poised, enthusiastic, and confident.	
	D. 2 Excessive use of distracting mannerisms such as "Ahs" and "Ohs" not evident.	
	D. 3 Instructor's voice quality, volume, and variation (pitch, rate, inflection) appropriate for group and classroom size.	
3. Varies activities with which the instructional stimuli are presented; e.g., lecturing, questioning, discussion, practice (daily).	F. 1 Training aids/instructional material/equipment listed in Plan of Instruction (POI) used by the instructor/students during lesson.	
	F. 2 Chalkboard and other visual aids used in an effective manner.	
4. Uses a mix of rewards and reinforcers (weekly, monthly).		
5. Varies types of questions and probes.	Q <u>convergent</u> <u>divergent</u> <u>to clarify</u> P <u>to solicit</u> <u>to redirect</u>	G. 1 Questions phrased clearly and to the point. G. 2 Questions appropriate for the lesson. G. 3 Variety of question types used. Types of question used adjusted to the situation.
6. Uses learner ideas and participation to foster lesson objectives when appropriate (weekly).		E. 2 Students involved in demonstration, if appropriate. H. 5 Student participation encouraged. I. 2 Group members encouraged to participate.

Table 9 (Continued)

Key behavior	Teaching effectiveness Indicators from research literature	Related items on Form 281
Task Orientation	<ol style="list-style-type: none"> Develops unit and lesson plans in accordance with text and curriculum guide. Handles administrative and clerical interruptions efficiently. Stops misbehavior with a minimum of disruption to the class. Generally, uses direct instruction strategies for teaching Type I behaviors and indirect instruction strategies for teaching Type II behaviors. Establishes end products (e.g., reviews, tests) that are clearly visible to students. 	<p>B. 2 All objectives covered. Lesson flowed smoothly from point to point.</p> <p>A. 3 Training materials, aids, and equipment available prior to the start of class.</p> <p>A. 4 Necessary materials distributed effectively.</p> <p>H. 1 Proper control of class maintained.</p> <p>J. 1 Progress check and performance/written tests administered properly.</p>
Engagement in Learning Process	<ol style="list-style-type: none"> Provides for guided practice. Provides correctives for guided practice in a non-evaluative atmosphere. Uses individualized or attention-getting strategies to promote interest among special types of learners, when appropriate. Uses meaningful verbal praise. Monitors seatwork by circulating and frequently checking progress. 	<p>G. 4 Student questions answered adequately.</p> <p>H. 2 Appropriate techniques used to assist and motivate students.</p> <p>J. 1 Progress check and performance/written tests administered properly.</p>

Table 9 (Concluded)

Key behavior	Teaching effectiveness indicators from research literature	Related items on Form 281
Moderate to High Rates of Success	<ol style="list-style-type: none"> 1. Unit and lesson organization reflects task-relevant prior learning. 2. Administers correctives immediately after initial response. 3. Divides lessons into small, easily digestible chunks. 4. Plans transitions to new content in small, easy-to-grasp steps. 5. Establishes momentum; e.g., pacing and intensity gradually build toward major milestones. 	<p>A. 1 Lesson plan current, personalized, and appropriate. Lesson plan signed by instructor's supervisor.</p> <p>K. 1 Thorough knowledge of subject matter demonstrated.</p> <p>A. 4 Student questions answered adequately.</p> <p>A. 1 Lesson plan current, personalized and appropriate. Lesson plan signed by instructor's supervisor.</p> <p>H. 3 Time managed appropriately. Lesson well paced.</p>

As can be noted, several of the above items are specific to the military classroom and, therefore, would not be expected to have counterparts in the teaching effectiveness research conducted in public school classrooms. Conversely, it can be noted from Table 9 that 9 of the 28--or approximately one-third of the teaching dimensions identified in the teaching effectiveness literature--do not directly correspond with any item on Form 281, suggesting the possibility of expanding the content of Form 281 to reflect these behaviors, where applicable to the ATC classroom.

Instructor Evaluation in the Army and Navy and Some Comparisons with Form 281

This section briefly reviews the instructor evaluation forms used by the Army and Navy.

Army. Similar to the Air Force, the Army uses a single-page, two-sided form for instructor evaluation. Five general teaching effectiveness factors are measured on this form, comprising a total of 25 items, as compared to 35 items on ATC 281. These items are distributed across five basic factors, as compared to 12 factors on ATC Form 281. These factors are:

- Introduction/Statement of Objectives (3 items)
- Lesson Presentation (6 items)
- Instructor Communication Skills (6 items)
- Instructor Preparation (7 items)
- Training Site Management (3 items)

Each of these factors is similar to one or more of the 12 general factors measured on ATC Form 281. Items on the Army form, however, employ a 7-point scale, as compared to a 4-point scale on ATC Form 281, and representing the following intervals.

- NR = Not Relevant
- 1 = Failing
- 2 = Poor
- 3 = Unsatisfactory
- 4 = Marginally Satisfactory
- 5 = Fully Satisfactory
- 6 = Outstanding

An additional difference between the Army form and ATC Form 281 is that each item on the Army form is assigned a factor weight ranging from 1 to 6, depending on its perceived importance to Army instruction. Hence, an item rated "6" and assigned a weight of "5" would be scored as "30." The maximum score attainable across all items, if each item were rated "6" and multiplied by its proper weight, is 600. This base is then adjusted downward for the number of nonrated items by multiplying the total number of nonrated items by 6 and subtracting this value from the base. This adjusted base is then divided into the weighted sum of all the items and multiplied by 100 to arrive at a percentage of the total points possible. The instructor is then assigned an overall rating based on the following ranges of percentages:

- 0 - 66 Below Average
- 67 - 83 Average
- 84 - 94 Superior
- 95 - 100 Exceptional

Space is provided at the end of each of the five sections for comments pertaining to that section. Space is also provided at the end of all sections for general comments. A critique and feedback session with the instructor is required no later than 48 hours after the assessment. A

matching of items on the Army instructor evaluation form with the teaching effectiveness indicators presented in Tables 1, 2, 3, 6, and 7 reveals that 64% of the items have some degree of correspondence with the teaching effectiveness indicators commonly referenced in the literature, as compared to 80% for ATC Form 281. As in the case of Form 281, 5 of the 25 items on the Army instructor evaluation form appear to be military-specific and, therefore, would not be expected to have counterparts in the teaching effectiveness research conducted in public school classrooms.

A copy of the Army evaluation form appears in Appendix D.

Navy. Although the Air Force and Army instructor evaluation procedures are somewhat similar in approach, the Navy instructor evaluation procedure differs considerably from both in scope and depth. The Navy approach to evaluation is referred to as the Course Evaluation System (CES), emphasizing its focus on the evaluation of the organization of the subject matter being taught (the course), as well as the presentation or delivery of the subject matter by an instructor. In the Navy system, the evaluation of the individual instructor is only a part of this more general evaluation system.

The CES is divided into three distinct parts, any one of which may be completed to the exclusion of the others. Also, due to the labor-intensive nature of the CES, use of the system is recommended only if a course or instructor is shown to be having "serious problems," such as negative feedback from superiors, high attrition, critical student comments, poor student performance on tests, negative instructor review, or an unusual number of setbacks. The three major components of the CES are:

Objective Classification

Objective and Test Adequacy and Consistency Evaluation

Presentation Evaluation

The first two components will be summarized briefly, as only the third component, Presentation Evaluation, involves the direct evaluation of instructors.

Objective Classification. For this operation, long-term or terminal course objectives are subdivided into enabling objectives and classified according to (a) whether the objective requires the student to remember or use information; (b) if "use," whether use should be aided (for example, with diagrams) or unaided; and (c) whether or not the objective requires the transfer of learning to some applied context. Therefore, each enabling objective for the course is simultaneously classified as remember or use, aided or unaided, and transfer or not transfer.

Objective and Test Adequacy and Consistency Evaluation. Next, test items from the course examination are matched to each enabling objective using a matrix-type format. In addition, a wide variety of additional information is collected and associated with each enabling objective including (a) the general task to which the objective applies; (b) whether supporting objectives are required; (c) whether the objective is "essential," "nice-to know," or "unnecessary"; (d) the training goal to which the objective applies; (e) whether the objective includes a statement of conditions, standards, and actions; (f) whether the test item matches the objective; (g) whether the test item matches the training goal; and (h) whether the test item is adequate. Preferably these evaluations are to be carried out by two or three subject-matter experts in the content area being evaluated.

Presentation Evaluation. This component of the CES applies to the evaluation of the course instructor delivering instruction in a classroom or laboratory setting. Even here, however, the evaluation is considerably broader than Air Force or Army procedures in that it first requires evaluating the instructor's performance with respect to its consistency with the objectives

identified in the previous components and the test items on the final course examination. Here, the instructor's lesson plans, student guides and classroom presentation--generally over repeated lessons--are studied to determine if the instruction including media, materials and handouts matches the objectives and test items. Each objective is marked complete, incomplete, or not applicable on each of the following dimensions: whether the objective was clearly stated, whether practice in remembering was provided, whether feedback was given, whether examples were provided, whether an opportunity to practice using the information was afforded, and whether feedback followed this practice opportunity.

Next, the presentation of each objective is rated for adequacy by indicating whether the practice, feedback, and examples provided for each objective were:

1. separate from other segments of the instruction
2. properly identified
3. clearly stated
4. accompanied with student help
5. job-oriented
6. sufficient in number
7. sufficient to detect student errors
8. provided in an easy to hard progression
9. provided in a way to encourage student involvement

From a review of the instructor's media, materials and handouts, and classroom observation, raters respond "Yes" or "No" to each dimension applicable to the practice, feedback and examples provided for a specific objective.

The third and final part of the Navy's instructor evaluation parallels ATC Form 281 and its Army equivalent. This phase of the evaluation employs an Instructional Effectiveness Checklist comprised of five general factors and 47 items. These general factors are:

Learning Orientation (7 items)
Instructor Behavior (19 items)
Media Materials (5 items)
Environment/Safety (9 items)
Student Behavior (7 items)

Each item is checked either "Yes" or "No." No space is explicitly provided for comments on the form, although the margins could conceivably be used for this purpose. Unlike the Air Force and Army forms, this form devotes separate factors to environment/safety and to student behavior. Additionally, the 7 items devoted to student behavior (e.g., Did it appear the students achieved the objective? Did it appear the students took an adequate amount of notes? How involved were the students in the learning process?) have no counterparts on the Air Force and Army forms.

A matching of items on the Instructional Effectiveness Checklist with the teaching effectiveness indicators identified in Tables 1, 2, 3, 6, and 7 reveals that 50% of the items have some degree of correspondence (as compared to 80% for the Air Force form and 64% for the Army form), after items pertaining to student behavior are eliminated (since the forms of the other services do not include this dimension). Also, about 10 of the 40 items included in this analysis appear to be military-specific and, therefore, inapplicable to the research in public school classrooms. Finally, it can be noted that, contrary to Air Force and Army procedures, the Navy's Instructional Effectiveness Checklist, as well as the other components of the Course Evaluation System, is recommended and one presumes largely used only when "serious problems" come to the foreground to justify the labor-intensive process involved. The regularity of its use and the context within which feedback is provided are unclear from currently available documentation. The Navy's Instructional Effectiveness Checklist appears in Appendix D.

V. SURVEY OF OBSERVATION AND MEASUREMENT FORMATS SUITABLE FOR FEEDBACK ENHANCEMENT

This section surveys several measurement formats typically used in recording effective teaching behaviors, such as those reviewed in Section II, with particular emphasis on the advantages and disadvantages of each format for providing instructional feedback. This section includes the identification of those individuals who might serve as data recorders/observers in the ATC instructional environment, along with a prioritization of formats using the criteria of acceptability to users, ease of implementation, cost, objectiveness, increased feedback, validity, and reliability. The measurement formats reviewed and compared in this section are:

- Time Interval Observation
- Likert Scales
- Semantic Differential Scales
- Guttman Scales
- Checklists

As the technical details of these formats are amply covered in textbooks on measurement and testing (see, for example, Kubiszyn & Borich, 1987), they will be only briefly summarized here for the purpose of comparing the unique strengths and weaknesses of each format for providing feedback to ATC instructors.

Time Interval Observation

One measurement tool that has been used in the evaluation of teaching is the observation of discrete categories of teaching behavior sampled over intervals of time. Two characteristics distinguish various observation instruments of this type: (a) the recording procedure, and (b) the item content.

Tools for observing continuing classroom events may employ either of two recording procedures--sign or category. A sign system records an event only once regardless of how often it occurs within a specified time period. The behavior is given a code (e.g., "/"), which merely indicates its presence or absence within a particular block of time. A category system, on the other hand, records a given instructor behavior each time it appears or every time it appears within a brief designated interval of time and, hence, provides a frequency count for the occurrence of specific behaviors, rather than a mere indication of their presence or absence. A frequency count may also be obtained using a modified sign system, called a rating instrument, which estimates the degree to which a particular behavior occurs. For example, instead of simply noting the presence or absence of a behavior, a rating instrument may indicate the relative

frequency at which the behavior occurs on, say, a 1-to-5 scale, with "5" indicating a high frequency of occurrence (e.g., "occurs often") and "1", a low frequency of occurrence (e.g., "never occurs").

Time interval observation systems can be further differentiated on the basis of their item content. Generally, observation instruments, whether of the category, sign, or rating variety, focus on either high- or low-inference behaviors. Those instruments which ask an observer to judge, for example, the general degree of a teacher's clarity, variety, or task orientation require high inference, since the item content (e.g., the word "clarity") does not specify discrete behaviors that must occur in order for a teacher to be considered clear, flexible, or task-oriented. In this instance, the item content is integrative and cumulative in nature, forcing the observer to make judgments about the behavior being observed over a relatively long period of observation. On the other hand, observation instruments that name specific behaviors to be recorded, such as "instructor asks questions" or "instructor uses examples," require less inference on the observer's part. Low-inference item content such as this generally reflects separate and distinct units of behaviors which are easy to observe over relatively brief spans of time, sometimes consisting of only a few seconds. One unique feature of low-inference observation systems is that, due to their discrete behavioral content and brief time intervals, they can be used to record the sequence or flow of classroom behaviors from instructor to student and from one interval to the next. It should be noted, however, that not all observation systems are either high- or low-inference. Some combine the two types of item content, while others require an intermediate level of inference from the observer.

Figures 1, 2 and 3 illustrate differences in the recording procedures and time intervals commonly used in the sign, category and rating types of observation systems.

Ten-minute intervals						Instructor practices
I	II	III	IV	V	VI	
						1. Reviews or summarizes.
						2. Gives examples.
						3. Uses advanced organizer.
						4. Asks questions, probes.
						5. Demonstrates.
						6. Gives directives.
						7. Provides feedback.

Figure 1. Sign System. (A check mark is given to the behavior that best characterizes the interval.)

Category	Instructor						Learner			Total
	1	2	3	4	5	6	7	8	9	
Reviews or summarizes										0
Uses examples										0
Provides advanced organizer	1			1						2
Asks question		2		1			12		1	16
Demonstrates	5	22	3							30
Gives directives			1	5		3		4	13	
Provides feedback correctives										0
Responds	1	7	4	4		14		1	31	
Initiates										0
Silence		2	1	1		2		3	9	
Total	0	0	2	16	30	13	0	31	0	101

Figure 2. Category System for Recording Sequential Events.
(A tally is made each time the behavior occurs.)

1. Enthusiasm: Enthusiastic - Apathetic	Enthusiastic	Neutral	Apathetic
	1	2	3
2. Classroom Environment: Supportive - Nonsupportive	Supportive	Moderate	Nonsupportive
	1	2	3
3. Task Orientation: Focused - Unfocused	Focused	Moderate	Unfocused
	1	2	3
4. Clarity: Clear - Unclear	Clear	Moderate	Unclear
	1	2	3
5. Structuring: Structural - Unstructural	Structural	Moderate	Unstructural
	1	2	3
6. Variety of Methods: High - Low	High	Moderate	Low
	1	2	3
7. Cognitive Variety: Varied - Unvaried	Varied	Moderate	Unvaried
	1	2	3
8. Amount of Criticism: High - Low	High	Moderate	Low
	1	2	3

Figure 3. Rating System after a 1-hour interval of observation.
(The appropriate degree of behavior is circled)

An important distinction among time interval observation instruments involves differences in length of the observation interval. An observation interval is the period of elapsed time allowed for a single discrete observation and for recording the results of that observation. Recording procedures can differ widely with respect to length of interval. The sign method--or the process of recording the presence or absence of a behavior within an interval of time--is the most common among observation systems. However, this simple notion can take on different qualities depending on the length of the interval of observation used. Sign systems have been constructed with relatively long time intervals (from about 15 minutes to an entire class period) as well as relatively short time intervals (from about 10 seconds to 1 minute). Generally, however, when a less dependent, less discrete and more general (higher-inference) behavior is the focus of interest, the more appropriate is a long time interval. Conversely, the more independent, more discrete and more specific (lower-inference) is the behavior of interest, the more appropriate is a short time interval. This reflects the notion that when the behavior of interest is occurring rapidly and independently of other behaviors, a short interval would be needed to capture frequent shifts in its status. When the behavior of interest occurs relatively infrequently and/or its recognition is affected by other events, a longer interval would be needed before its presence or absence could be detected.

Another measurement issue related to the length of an interval is the extent to which the observer processes and integrates information throughout the interval or simply records the behavior present at the exact moment of observation. Long interval systems, say on the order of 10 minutes or longer, generally attempt to depict the behavior that best characterizes the entire interval. On the other hand, short interval systems, say on the order of 1 minute or less, tend to force observers, whether intentionally or not, to use a "snapshot" technique requiring little or no processing and integrating of information beyond that needed to recognize a behavior at an exact moment of time. These two types of intervals can provide different results, since in the former case the accuracy of the data depends, in part, on the abilities of the observers to process and to integrate information during the observation interval, whereas in the latter case, the accuracy of the data depends on the ability of the observers to simply recognize the behavior at a given instant.

Summated Ratings

Summated ratings can be divided into Likert and Semantic Differential Scales.

Likert Scales. The Likert scaling technique requires a large number of items that describe teaching behaviors, each yielding a high score for a favorable rating on a behavior and a lower score for a less favorable rating. The rater reacts to items on a 5-point response continuum which reflects either the quality of behavior or the frequency at which it was perceived to occur. The Likert procedure customarily yields scales with moderate to high reliability. Validity, however, can vary, due to the following considerations:

- Generally, no attempt is made in the construction of a Likert scale to ensure equal distances between units (e.g., between "very often" and "fairly often," or between "always relevant" and "most relevant"). Therefore, increments of change may have different meanings at different parts of the scale.
- The unidimensionality of the scale (i.e., the extent to which it measures a single, distinct behavior) must be inferred from high correlations between item and total scores. Low item/total correlations indicate that the construct is too multifaceted and factorially complex to allow simple and direct interpretation.
- Likert scores generally are interpreted according to a distribution of sample scores, and an individual instructor's score has meaning only in relation to the scores of other instructors who have been similarly rated. In other words, these scales provide ratings that are relative to others being rated and not to some absolute standard or criterion.

A typical Likert scale item might appear as:

Instructor uses advanced organizers

_____ very often
_____ fairly often
_____ sometimes
_____ fairly rarely
_____ very rarely

Semantic Differential Scales. The semantic differential scale is another summated rating method used to cumulatively record the quality or frequency of teaching behaviors. It requires the rater to judge the instructor's performance on a series of 7-point bipolar scales. The rater checks the appropriate space, indicating both the direction and intensity of the judgment. Scores are derived by assigning numbers (i.e., 1, 2, 3, 4, 5, 6, and 7) to each position on the rating scale. Since the semantic differential and Likert scales are similar, the cautions noted above also apply. The semantic differential type does not necessarily exhibit equal intervals between scale points; the unidimensionality of the concept being measured may vary from one scale to another (particularly when bipolar responses are not exact opposites); and scores are interpreted relative to the rated performance of others. In practice, differences between Likert and semantic differential scales are minor and generally are related to the use of 5- or 7-point response formats. The similarity of these procedures is often reflected by high or moderate correlations between the two when they are used to measure the same behavior. A portion of a typical semantic differential scale might appear as:

Today's Lesson

Structured	_____	_____	_____	_____	_____	_____	_____	Unstructured
Clear	_____	_____	_____	_____	_____	_____	_____	Unclear
Focused	_____	_____	_____	_____	_____	_____	_____	Unfocused

Scalogram Analysis (the Guttman Scale)

Another method that is used to record judgment of teaching performance is the Guttman scale. This method is based on the idea that some behaviors can be arranged hierarchically so that an instructor who manifests a particular behavior at some level of complexity may be assumed to possess all other, related behaviors of lesser complexity. For example, if an instructor correctly uses visuals in the course of teaching, it may be assumed that the instructor also can correctly show to another instructor how visuals can be used in a written lesson plan and can recall three of the most prominent uses of visuals. When such an arrangement is found to be valid, the behaviors are said to be scalable.

In developing the Guttman scale, items are formulated and arranged in a most to least complex--or hierarchical--order. These items are then pilot-tested on a group of instructors, whose response patterns are then analyzed to determine whether or not the items are scalable. If items require only agreement or disagreement (i.e., an indication of the presence or absence of a behavior), only $n + 1$ of these patterns can be obtained. The relatively low frequency of deviant patterns allows the computation of what is called a coefficient of reproducibility (R). R is equal to the proportion of responses that can be correctly reproduced from the knowledge of an instructor's score. The extent to which such inferences can be made depends upon the level of the coefficient of reproducibility. This value represents a measure of the unidimensionality of the scale and is an index of the scale's validity.

Like the Likert and semantic differential scales, the Guttman scale makes no attempt to ensure equal units between items. However, unlike the Likert and semantic differential types, the Guttman scale need not be interpreted relative to the ratings of other instructors, since its items represent specific discrete behaviors, the presence or absence of which can form the basis of an absolute as well as a relative judgment. Generally, this is considered a desirable characteristic for measuring teaching competence. A typical Guttman Scale appears below.

- 0 Instructor can elicit increased pupil achievement with use of visuals.
- 0 Instructor can correctly use visuals in the course of teaching lesson content.
- 0 Instructor can correctly show how visuals can be used in a written lesson plan or simulated environment.
- 0 Instructor can orally explain proper use of visuals.

0	0
0	0
0	0
0	0
Expected pattern	Deviant pattern

The expected pattern indicates one of several response arrangements wherein the item would be considered scalable (high coefficient of reproducibility, if found to occur over a large number of individuals). The deviant pattern indicates one of several response arrangements wherein the item would be considered unscalable (low coefficient of reproducibility) and in need of revision and further pilot-testing.

Checklists

When a behavior cannot be easily rated on a continuum of values, a simple indication of its presence or absence is used. If an observer is unable to make fine gradations in judging the quality or frequency of behavior, a simple yes-no, observed-unobserved, or present-absent format is sometimes used. Since checklists record only the presence or absence of behaviors, they assume that the rater has had ample opportunity to observe the behavior in question and the instructor has had ample opportunity and time to display the behavior. However, this assumption may at times be unwarranted. When the checklist data indicate the absence of a particular behavior, it should be determined whether this reflects a true absence or simply a lack of opportunity to observe the behavior. The latter situation may occur, for example, when the instructor's objectives are unrelated to or incompatible with the particular behavior in question, or when the rater has visited the classroom too infrequently or for too short a period to have had an opportunity to observe the behavior. In order for the rater to distinguish the absence of an event from inadequate opportunity to observe the event checklists should provide two separate response alternatives: (a) opportunity to observe the event, and (b) presence of the event. The rater would check the first alternative whenever a behavior on the checklist can be observed, given existing classroom conditions. The "true" presence or absence of a behavior would then be recorded by checking or failing to check the second alternative, but only after a check has indicated an opportunity to actually observe this behavior exists, given the existing classroom conditions. This format is illustrated below.

Opportunity	Present	
0	0	Instructor uses advanced organizers
0	0	Instructor reviews and summarizes
0	0	Instructor provides examples and demonstrations

The following summary lists the advantages and disadvantages of each of the above measurement/observation formats.

Time Interval Observation

Advantages

1. Can provide a large representative sampling of behavior.
2. Can record behavioral sequence of instructor-student dialogue.
3. Can provide a detailed longitudinal record, week-by-week or month-by-month, of an instructor's performance.

Disadvantages

1. Provides large amounts of data that must be reduced through statistical analysis.
2. Requires the training of observers to assure reliable coding.
3. Can be time-consuming to process and interpret if large amounts of data are collected.

Rating Scales

Advantages

1. Directs observation toward specific and clearly defined aspects of behavior.
2. Provides a common frame of reference for comparing all individuals on the same set of characteristics.

Disadvantages

1. Results in a tendency to rate all individuals at approximately the same position on the scale.
2. Can allow a rater's general impression of the person to influence how that person is rated on individual characteristics.
3. Usually allows for relative judgments only; that is, allows for comparisons among individuals but not absolute judgments that a particular behavior did or did not occur.

Checklist

Advantages

1. Is useful in evaluating those performance skills that can be divided into a series of clearly defined, specific actions.
2. Can provide absolute judgments of the presence or absence of specific learning outcomes.

Disadvantages

1. Is not useful in recording general impressions of behaviors that are customarily seen as continuous; for example, an instructor's enthusiasm, experience or attitude.
2. Can provide inaccurate data when the opportunity to observe the behavior is not recorded at the time of scoring.

Guttman Scale

Advantages

1. Due to its hierarchical construction, can provide information as to whether different behaviors have been attained without actually having observed all of them.
2. Provides relative as well as absolute ratings.

Disadvantages

1. Requires considerable knowledge of the behaviors being measured in order to arrange them into a hierarchy.
2. Is applicable only to those behaviors that can be related to one another in a hierarchical, stair-step fashion.
3. May require extensive pilot-testing

Table 10 presents these measurement systems in relation to five different groups of individuals who could potentially serve as observers and/or raters in an enhanced evaluation and feedback system. These groups are: the designated instructional supervisor within a particular ATC content domain, the senior instructor, peer instructors, students and the instructor himself/herself using videotaped episodes of his/her own teaching. Theoretically it would be possible for each of these groups or combinations of them to provide data for instructor feedback using one or more of the measurement formats listed. There are, however, tradeoffs that would need to be considered in selecting any measurement system-data provider combination.

Table 10. Measurement Systems by Possible Data Providers

Supervisor	Senior Instructor	Peer Instructors	Students	Videotape Self Evaluation
Time				
Interval				
Observation				
Ratings				
Guttman Scale				
Checklist				

Using the criteria of acceptability to the user, ease of implementation, cost, objectivity, and increased feedback and validity/reliability, estimates of these tradeoffs can be made. First, it can be noted that greater differences in tradeoffs are likely to occur from one measurement system to another (down the columns of Table 10) than from one data provider group to another (across the rows). This is to say that, for example, acceptability to users or cost is likely to vary considerably more across measurement systems than, say, from supervisors to senior instructors. Although some variation among these groups in perceived acceptability, ease of use, cost, objectivity, and increased feedback and validity/reliability may occur, these sources may be considered moderately or highly correlated with respect to the judgment criteria being used. Therefore, the following description summarizes the most salient differences across measurement systems, focusing on data provider groups only when there is likely to be a significant departure in acceptability, ease of use, cost, objectivity, feedback, and validity/reliability among these groups. Although somewhat speculative, the following section points out some of the advantages and limitations of these measurement systems for enhanced instructor feedback in the ATC instructional environment.

Some Advantages and Limitations of Time Interval Observation

Time interval observation often is considered among the most expensive methods of providing observation and feedback data. For this reason, its use has been limited largely to classroom research on teaching effectiveness. Its cost derives largely from the extensive training and coding manuals that are often necessary to secure high levels of reliability by those doing the coding. This also reduces its ease of use in comparison of other measurement systems inasmuch as training workshops and material (manual) development is usually mandatory before coders can become proficient at perceiving and categorizing classroom behavior at a relatively rapid rate (e.g., 10-, 20- or 30-second intervals). This problem becomes less severe when a sign system (as opposed to a category system) is used and time intervals may be extended to 10 or 15 minutes. However, in this case, the system functions much like a simple checklist. When sufficient resources have been devoted to coder training and manual development, the objectivity, amount of feedback and reliability of a time interval system is generally considered to be high. However, little is known about the validity of these systems except that the data provided by them sometimes vary depending on the length of the time sampling interval chosen, which is often arbitrary. Finally, it is important to note that time interval observation, especially when the length of the interval is small can produce voluminous frequency counts of individual behaviors, and can easily befuddle and overwhelm a naive user. This is yet another reason why time interval observation has found greater acceptance among classroom researchers than practitioners who may be expected to use or feed back the results of the measurement system in a real-time environment. There is likely to be little difference among these evaluative criteria across data provider groups, except to note that this measurement system would not be suitable for use by students due to its time and technical demands. It would, on the other hand, be more suitable for use by the instructors themselves using videotaped segments of their own or their peers' teaching behaviors. In summary, a time interval observation system can provide extensive amounts of specific, discrete data about an instructor's use of specific behaviors via frequency counts. On the other hand, a time interval observation system is a form of measuring classroom behavior that may be cumbersome to use and that is likely to place considerable technical and time demands on its users. Its use seems primarily suited for classroom research and only secondarily for instructional feedback.

Summated Ratings (Likert and Semantic Differential)

Likert and semantic differential scale-type ratings are among the most popular forms of instructor assessment. The quality of feedback instructors typically receive from these types of

ratings, however, has been known to vary depending on the thought, effort, and knowledge of the instructional process that has guided development of the form itself. Generally, acceptability among users is moderate to high with these types of systems, since most instructors have long been accustomed to using them and being evaluated by them. Their acceptability, however, tends to be higher for evaluation purposes than for feedback, in that most summated types of ratings tend to provide only relative and not absolute assessments of a particular teaching behavior. That is, they most typically indicate that more or less of a given behavior is being observed relative to others who are being rated, but not whether the behavior is at some predesignated standard of acceptable performance. This means that moderate or even sometimes high ratings can mask poor performance when the behavior observed for an individual is actually only slightly better than the typical but poor behavior of others who are rated with the same scale. This tends to make summated ratings more useful for making comparisons across individuals than for providing explicit feedback to a single individual as to which effective teaching behaviors were or were not present. All summated ratings need not be only relative assessments, however. Constructing scale items that identify explicit teaching behaviors derived from the research literature on effective teaching and reducing scale intervals to only those variations of the behavior that can be reliably detected are ways in which summative ratings have been made more useful for providing instructional feedback.

Finally, it can be noted that since summated ratings often take the form of relative assessments, they are known to be influenced by who is doing the rating and who is being rated. That is, the rater's general impression of the person may influence how the subject is rated and, conversely, the characteristics of the rater and his/her belief or value system can, to lesser or greater degrees, influence the ratings. This problem is compounded by the tendency to rate most or all individuals at approximately the same position on the scale, due to the rater's belief in how the data may be used to help or hinder particular individuals. Hence, the lack of variation in summated ratings can prevent the provision of meaningful feedback to the person being rated.

In summary, summated ratings (i.e., Likert and semantic differential scales) are easy to use, are relatively inexpensive to construct and are a familiar and readily acceptable form of measurement to most users. Summated ratings are vague or abstract allowing the opportunity for personal beliefs and values to influence the rating. Also, since most summated ratings are assessments of the degree to which a particular behavior occurs relative to others who are being rated, as opposed to identifying the absolute presence or absence of a particular teaching practice, they tend to be more useful for evaluation than for feedback. It can be noted, however, that with proper pilot-testing, summated ratings frequently produce moderate to high reliabilities and modest predictive validity. Differences among these criteria are not likely to occur across the data provider groups identified in Table 10.

Some Advantages and Limitations of the Checklist

Checklists typically have been the easiest to use and most acceptable to users. Their ease of use derives largely from their binary (present, absent) nature and the increased certitude raters generally experience in having to make distinctions of only a "yes, no," "present, absent," "observed, unobserved" variety, when item content is sufficiently discrete as to lend itself to such all-or-none judgments. This has created both respect for and criticism of the checklist technique among various user groups, depending on the nature of the item content and the purpose for which the data are to be used. Generally, when item content is more readily perceived as a continuously occurring behavior in real life and when data are to be used for evaluation purposes, the checklist format has not been the method of choice. Individuals rarely prefer to rate or see themselves rated on an all-or-none basis when natural variation normally occurs and can be pointed to in a real-time environment. In these cases, the checklist suffers from overly reducing the complexity of performing a teaching act by making trivial or simplistic

distinctions that do not normally occur in the classroom. On the other hand, when certain behaviors do occur as discrete events, and degrees of variation are difficult to detect reliably, the measurement of that behavior is more suited to a checklist. Furthermore, when decision making is reduced to the presence or absence of a behavior, feedback becomes absolute, as opposed to relative, enhancing the diagnostic value of the checklist data. That is, recommendations become explicit and instead of informing the instructor that relative to the others he/she is higher or lower on that behavior, the instructor is told that a specific competency has or has not been observed.

For these reasons, if used with the proper behavioral content and for feedback or diagnostic purposes, the checklist is generally acceptable, easy to use, relatively cost efficient, and capable of providing useful feedback. On the other hand, its objectivity and validity depend heavily on the relationship between the behavior being observed and some externally anchored measure of teaching effectiveness, such as student achievement.

The research evidence does provide some links between the presence and absence of individual teaching behaviors (e.g., use of advance organizers, questioning, review and summarization) and student achievement, as indicated in the review of the literature in Section II of this paper, some of which tends to support a checklist format. For these types of behaviors it would be unlikely or difficult to observe extensive variation, or such variation would not be deemed useful for diagnosis and feedback. It can be presumed, however, that not every effective teaching behavior would conform to this characteristic. Finally, it can be noted that due to the restricted response range for checklists, reliabilities can be among the highest among the measurement systems being reviewed here, when the content being measured is discrete. In at least some instances, summated ratings have shown less than satisfactory reliability because item scales were asking for more discrimination, however theoretically justified, than observers could practically detect in a real-time environment. With respect to data provider groups, it would appear that the checklist would be equally applicable to all groups, with the exception that checklists are sufficiently simple to use in real time that they would not require the use of videotapes nor would their reliability likely be enhanced by them.

Some Advantages and Limitations of the Guttman Scale

The Guttman scale is considered among the most psychometrically sophisticated teacher evaluation tools. Its sophistication, however, comes at considerable cost in both the expertise needed to devise hierarchical scales and the pilot-testing needed to obtain sufficient degrees of scalability. Perhaps foremost among its strengths, however, is its capacity to provide data for the diagnosis of teacher competencies, assuming the competencies are scalable. Rather than providing relative ratings, as do many measurement tools, the Guttman scale requires the identification of teaching practices for which discrete classroom operations can be observed along some type of continua (e.g., simple to complex). The idea behind the Guttman scale is that ultimate competency in an individual behavior can be observed in a stair-step or building-block fashion, with each earlier step representing a necessary requisite part of each subsequent step. In this manner, knowledge as to where an individual instructor falls on any step automatically reveals what other steps have and have not been achieved. Hence, with the observation of a single operation, information about all other operations pertaining to that behavior is simultaneously obtained.

As was noted, while the amount of feedback and diagnostic information is high with a Guttman scale, other practical considerations often temper its use. Some training of observers is generally considered important for obtaining reliable scales, as is some undergirding theory or rationale that can reveal the individual steps or blocks that might go together to comprise an item scale. Without a keen understanding of the teaching practice being assessed, pilot-testing

is reduced to an unsystematic process of trial and error which attempts to find the scalable components of a behavior. Also, it is safe to presume that not all effective teaching behaviors will have scalable components, in which case the method of choice would be a checklist or summated rating. In addition, acceptability and ease of use of the Guttman scale may be complicated by the fact that many practitioners are unfamiliar with its use. This may affect its objectivity and reliability unless some type of orientation and/or training is given to users. As with all measurement systems, the validity of the Guttman scale rests on its capacity to predict some subsequent index of teaching effectiveness. Generally speaking, properly scaled Guttman items have produced relationships to student achievement that have been among the highest of any measurement system. Due to the technical demands and content expertise required of the Guttman scale, however, its use has been limited to well-trained evaluators.

The following table ranks the four measurement systems according to the criteria of acceptability to users, ease of use, cost, objectivity, feedback, and validity/reliability.

Table 11. Ranking of Each of Four Measurement System Against the Six Criteria

Rank Acceptability	Ease	Cost	Objectivity	Feedback	Validity/Reliability
1. Ratings	Checklist	Checklist	Time Interval Observation	Time Interval Observation	Checklist
2. Checklist	Ratings	Ratings	Checklist	Guttman	Ratings
3. Guttman	Guttman	Time Interval Observation	Guttman	Checklist	Time Interval Observation
4. Time Interval Observation	Time Interval Observation	Guttman	Ratings	Ratings	Guttman

From the above table, it can be noted that the number of criteria for which a particular measurement system appears in the top 2 ranks is as follows:

System	Number of criteria for which system is Ranked 1 or 2
Checklist	5
Ratings	4
Time Interval Observation	2
Guttman	1

Overall, checklists and/or summated ratings tend to have the most potential for feedback enhancement. Given that each of these tools has its own measurement limitations and should not be judged more or less appropriate apart from the particular teaching behavior it purports to measure, some combination of summated ratings and checklist would seem more prudent than the exclusive use of a single measurement format.

VI. SUMMARY AND RECOMMENDATIONS

This section summarizes the results of this investigation and provides recommendations for revising and/or developing procedures for enhancing ATC instructor feedback.

Summary

Specifically, the goals of this project were:

1. To survey the state-of-the-art indicators of teaching effectiveness reported in the research literature.
2. To describe the implementation of teaching effectiveness indicators and related procedures being used to evaluate teachers in the public schools.
3. To provide an overview of the ATC instructor evaluation procedures and similar procedures in use by the Army and Navy.
4. To survey and prioritize observation and measurement formats suitable for feedback enhancement in the ATC instructional environment.

The purpose of investigating these four areas was to present a comprehensive set of data by which the Air Force could determine the feasibility and desirability of revising and/or extending existing ATC instructor evaluation procedures to enhance feedback to the instructor. The data, their implications, and the recommendations resulting from this report together represent a basis for deciding whether additional research and development work would benefit this feedback enhancement goal.

Phase 1 of this project surveyed the state-of-the-art indicators of teaching effectiveness reported in the research literature. In this phase a comprehensive review of the results of classroom research on effective teaching from 1960 to the present was undertaken. Five general dimensions of teaching effectiveness consisting of 28 distinct teaching behaviors were identified as having been consistently associated with increases in student achievement on standardized and classroom tests of academic performance. These 28 behaviors, along with rationale for their effectiveness, were provided in Section II of this paper.

Phase 2 of the investigation, then, turned to the extent to which the public schools were actually using these and/or other indicators of teaching effectiveness to evaluate and provide feedback to classroom instructors. It was found that the majority of states require a performance appraisal and feedback system of individual school districts using at least some research-based teaching effectiveness indicators similar to or identical with those reviewed in Section II. In one example, utilizing the performance appraisal system required by the State of Texas, 79% of the 28 effective teaching behaviors identified from the research literature had counterparts on that State's appraisal instrument. Typically, state appraisal instruments often include other, professionally based teaching behaviors in addition to those found in the research literature.

Phase 3 of this investigation examined the current ATC evaluation procedures for their feedback potential, with specific reference to Form 281 and Form 736, the student critique. It then compared these procedures with those currently in use by the Army and Navy. For this task the contractor observed classroom, laboratory and field instruction across 18 different specialties at four technical training centers (Lackland, Keesler, Sheppard, and Lowry AFBs). After each observation, the instructor and one or more instructor supervisors were interviewed as to their opinions about the capacity of Forms 281 and 736 to provide useful feedback and other topics related to feedback enhancement. In a separate analysis, it was determined that 80% of the

teaching behaviors on ATC Form 281 were found to overlap those found in the research literature reviewed in Section II. In comparison, 64% of the behaviors on the Army instructor appraisal form and 50% of the behaviors on the Navy appraisal form had counterparts in the research literature. The Air Force procedures for evaluating instructor effectiveness appeared sufficiently inclusive as to incorporate the most significant elements of evaluation in use by the other services. Additionally, this phase of the investigation found that instructors and supervisors believed that:

1. Several of the items on Form 281 were either redundant or highly correlated.
2. Some items on Form 281 were not under the control of the instructor.
3. Some items would be more appropriately scaled with a 2-point format (checklist) than a 4-point format.
4. Some areas important to an instructor's effectiveness (e.g., pertaining to the cognitive organization of the instruction and providing impromptu examples) were absent and some (e.g., pertaining to public speaking and communication) were overemphasized on Form 281.
5. In the absence of more explicit definitions, some items on Form 281 were ambiguous and subjective.
6. Form 281 is significantly less suitable and efficient for evaluation and feedback in a laboratory or field class than in a formal stand-up classroom lecture.

Phase 4 of the investigation surveyed the advantages and disadvantages of several types of measurement formats that might be used to provide additional feedback to ATC instructors. The measurement formats reviewed were:

1. Time interval sampling
2. Likert scales
3. Semantic differential scales
4. Checklists
5. Guttman scales

Each format was reviewed for its acceptability to users, ease of implementation, cost, objectivity, and increases to feedback, validity and reliability (see Table 11, p. 63). In addition, a list of possible ATC data providers consisting of supervisor, senior instructor, peer instructors, students, and videotape self evaluation were discussed in relation to the practicality of these individuals participating in the feedback enhancement process using these various measurement formats. In addition to prioritized rankings of each measurement format under each of the above six criteria (acceptability, ease, cost, objectivity, feedback, validity and reliability), the results indicated that checklists and Likert scales would be the most suitable and practical for a feedback enhancement system.

Recommendations

Recommendations pertain to the following six areas of ATC instructor evaluation and feedback:

1. ATC Form 281

2. Supervisor's training course
3. Evaluation of laboratory and field-based instruction
4. Peer evaluation of instructors
5. Student evaluation of instructors
6. Technical training instructor course

Specific development initiatives will be suggested in each of these areas, along with an approximate estimate of the number of man-days required to complete each development task, exclusive of pilot-testing and validation.

ATC Form 281

ATC Form 281 is the major source of instructor feedback currently used by the Air Force. Although most instructors and instructor supervisors agree that Form 281 appears adequate in its present form for evaluation purposes, the consensus is less strong that Form 281 comprises an adequate instrument for feedback to the instructor as to his or her teaching effectiveness. Problems pertaining to the feedback purpose of Form 281 which were mentioned during interviews with instructors and instructor supervisors were:

1. Some items on Form 281 may not always be under the control of the instructor. Some Form 281 items frequently mentioned were: A2 "Classroom neat and orderly, seating arrangement appropriate. Items identified needing repair, such as burned out lights"; and I1 "Students in compliance with appropriate dress and appearance standards." Instructors who rotate from classroom to classroom due to space and scheduling limitations (about 50% of the ATC instructors sampled) are quick to point out the irrelevance of Item A2 for feedback (and evaluation).
2. Some ATC Form 281 items appear redundant or highly correlated. Some instructor supervisors, for example, reported having difficulty distinguishing among such items as "Student participation encouraged" (H5) and "Group members encouraged to participate" (I2). Although the latter involves groups, the latter subsumes the former, thereby providing a lack of distinctive information across these items. In addition, some supervisors report that item G3 ("Variety of question types used. Types of questions used adjusted to situation.") is another form of encouraging student participation and, therefore, highly related to items H5 and I2, sometimes making meaningful distinctions among these items difficult.
3. Some ATC Form 281 items appear more suited to a present/absent scale than to the currently used 4-point rating system. For example, some instructor supervisors see no distinction between a rating of "Satisfactory" and one of "Outstanding" for items which primarily are intended to record whether a standard or regulation has been met or not met. Cited by instructor supervisors were items such as A1 ("Lesson plan current, personalized, and appropriate. Lesson plan signed by instructor's supervisor.") and A3 ("Training materials, aids, and equipment available prior to the start of the class.") Instructor supervisors frequently reported answering these types of questions as "Yes" or "No," thereby arbitrarily rating the instructor either as "Outstanding" or "Satisfactory," when a "Yes" response was in order. Some supervisors give only a "Satisfactory" rating when performance meets the implied standard, while other supervisors customarily give an "Outstanding" rating when the condition has been met, leading to inconsistencies across supervisor ratings for these types of items.

4. Some instructors and supervisors reported that ATC Form 281 tends to emphasize some of the most easily observable characteristics of presenting (e.g., eye contact, gestures, enthusiasm) and communicating (e.g., enunciation, mannerisms, voice quality) at the expense of other qualities which may have as much in common with effective teaching as good speaking. Although every instructor and supervisor interviewed reported on the close and necessary relationship between these two areas, some considered the absence or underemphasis of items such as the ones pertaining to (a) the accuracy of the content being presented, (b) the timeliness or recency of examples used to illustrate major points, (c) the ability of the instructor to use unexpected opportunities (such as particularly adept and unusual student response) to spontaneously foster lesson objectives, and (d) the use of advanced organizers to cognitively "frame" the lesson, as limitations of the usefulness of Form 281 for instructor feedback.

5. It was noted by both instructors and supervisors that few guidelines exist or have been communicated as to the extent and nature of the comments that should be provided on Form 281. Although any item receiving a rating of "Outstanding" or "Needs Improvement" must receive a comment in the right-hand margin, no specific guidelines or examples exist to encourage the feedback purpose on this portion of Form 281. As a result, feedback in the form of comments on Form 281 varies considerably across instructors, courses and technical training centers. A perusal of a sample of Forms 281 indicated that some supervisors comment on items rated "Outstanding" and "Needs Improvement" as specified in the regulations but seldom or never on items rated "Satisfactory." Other supervisors seem to prefer to comment extensively on many items in some areas of Form 281 but not at all on items in other areas. Generally, the instructors find the written comments helpful, but sometimes inadequate explanation as to the precise reason why a behavior was rated the way that it was. Often these comments are simply congratulatory if a behavior was well executed (but do not say what made it exemplary) or simply a call for improvement when the behavior was inadequate (but do not indicate how to improve it).

The above observations lead to the following suggestions for revising ATC Form 281 to enhance its instructional feedback characteristics:

1. Rework items on Form 281 to reflect only those responsibilities that are typically under the control of the instructor. Delete references to student behavior that may be only indirectly or tangentially related to effective classroom instruction and, therefore, not typically under the control of the instructor.

2. Partition (e.g., adherence to standards and regulations) Form 281 items into those most suitable to a checklist and those for which reliable discriminations can be made on the present 4-point scale. Reorganize Form 281 accordingly.

3. Replace redundant items on Form 281 with additional items that reflect the technical accuracy (e.g., K1 "Thorough knowledge of subject matter demonstrated" might be replaced with distinctions such as knowledge of mechanics, of concepts, of applications, of rules and regulations, etc.), timeliness, spontaneity and cognitive organization of the instructor's lesson. Add any behaviors from the effective teaching literature that are consistent with the 12 general dimensions which currently appear on Form 281.

4. Provide written guidelines for and realistic examples of comments reflecting diagnostic feedback and suggested sources of remediation for each item on Form 281. Make this written documentation available through the faculty development office of each Wing and the supervisor's training course.

The above revisions could be first pilot-tested as a new peer evaluation form, to be used on a voluntary basis over the course of 12 months, the data from which could then be used for evaluating any increase in feedback yielded by these changes. It is estimated that approximately 60 man-days would be required for completing these changes, exclusive of pilot-testing.

Supervisor Training Course

The supervisor training course is designed to acquaint master instructors with the supervisory and administrative duties they are about to assume as instructor supervisors. Although the composition of the course and areas of emphasis within it tend to vary slightly to reflect the needs specific to the technical training center at which it is taught, curriculum for the course specifies that an introduction to the use of ATC Form 281 must be taught. Generally, this introduction consists of a 10-minute film illustrating an Air Force instructor presenting a formal classroom lecture in which several ineffective teaching behaviors (e.g., failing to engage a student with a follow-up probe after a correct but hesitant response) can be observed. Although the film emphasizes "problem" behaviors, it also offers the opportunity to discuss, if not observe, effective as well as ineffective teaching behavior. Generally, the film is used to introduce Form 281, after which the items on a Form 281 relevant to the film are completed by each member of the class and then discussed. Although estimates vary from base to base, the time devoted to the use of Form 281 in the instructor supervisor's course is about 1 to 3 hours.

Recent graduates of the supervisor training course interviewed tended to report that instruction in the use of ATC Form 281 comprises an acquaintance with the rules and regulations pertaining to when and how to complete it. Generally, however, it was reported that this instruction excluded guidelines for completing the comment portion of Form 281; instruction on how to observe, mentally take notes of and record ongoing behavior; and instruction on how to conduct the post-observation briefing in which specific strengths and weaknesses should be diagnosed and suggestions for remediation provided.

Both instructors and supervisors reported considerable latitude in how ATC Form 281 was being used to provide instructor feedback. Therefore, the following recommendations are made:

1. The time devoted to ATC Form 281 in the instructor supervisor course should be upgraded to a highly organized simulation-intensive block of instruction.
2. Brief training films simulating effective and ineffective teaching behaviors should be produced for each of the 12 instructional areas on ATC Form 281. Each file would focus specifically on a designated area and its subcomponents, as noted on ATC Form 281. These areas are:

Preparation
Content/structure
Presentation skills
Communication skills
Demonstration skills
Use of training aids/materials
Question/answer techniques
Management
Supervising group activity
Measurement
Personal Qualities
Other

Supervisory trainees would be expected to learn how to accurately observe and record on ATC Form 281 effective and ineffective teaching behaviors in each of these areas.

1. Examples of diagnostically relevant and irrelevant comments and sources of remediation pertaining to each item on ATC Form 281 should be provided during the instructor supervisor course and made available to all in-service supervisors.

2. A role-playing performance exercise should be built into the supervisor training course in which each student conducts a post-observation briefing with a peer, using comments from a completed ATC Form 281. Instructional emphasis should be placed on the diagnostic accuracy of the comments made, their potential for constructive feedback to the instructor, and the specificity of the suggestions made for remediation. The exercise should be graded and an appropriate behavioral checklist prepared for student evaluation.

3. A manual to accompany ATC Form 281 should be prepared which defines, provides examples of and explains the rationale for each of the behaviors on Form 281. The purpose of the manual would be to standardize the meanings of each of the terms used on the form, providing a common reference and interpretation for all instructor supervisors. This product would extend but also considerably supersede in scope and update in accuracy ATCR 52-8, attachment 5, "Standards for Classroom Instruction to Assist Instructors in Using ATC Form 281."

Approximately 75 man-days are estimated for completing these recommended actions, exclusive of pilot-testing and media production.

Evaluation of Laboratory- and Field-Based Instruction

Although ATC Form 281 tends to emphasize formal classroom instruction, a large percentage of Air Force technical training is conducted in the laboratory or field. Approximately 50% of the instruction observed at Lackland (e.g., security police specialist course), Keesler (e.g., air traffic control specialist course), Sheppard (e.g., electrician specialist course), and Lowry (e.g., F-15 aircraft armament systems specialist course) depended heavily on laboratory, field and simulation experience. Instructors and instructor supervisors reported that, in these instructional settings, ATC Form 281 tends to be less useful as a feedback device than when instruction takes place in the formal classroom. Yet some instructor supervisors must use ATC Form 281 in these settings in order to meet the quarterly and 30-day evaluation requirements for their new (1 to 3 months) and relatively new (3 to 12 months) instructors.

For most supervisors this transfer to laboratory or field use is made by checking "NA" (Not Applicable) for items that are irrelevant. Most supervisors report trying to avoid using Form 281 in the field or laboratory when possible, but other supervisors report extensive use of Form 281 in these settings when 70% or 80% of their course is conducted on the flightline, in the hangar, or field shop. Supervisors report several problems when Forms 281 are used exclusively during formal classroom instruction in a course that is primarily field- or laboratory-dominated. When Forms 281 are used exclusively during formal classroom instruction in a field- or laboratory-dominated course, feedback specific to the field setting does not occur, even though a large percentage of the instruction takes place in these settings. Conversely, when Forms 281 must be used consistently in a field setting, sometimes a large number of items must be marked "NA" or else be interpreted broadly to fit the field setting, leaving some behaviors unevaluated or interpreted to mean something different for a particular supervisor at a particular training center. Therefore, the following recommendation is made:

1. Each technical area of training at the Wing, specialty or course level should be encouraged to develop a field- or laboratory-based counterpart to ATC Form 281, when applicable, which uniquely captures the performance, demonstration, and activity skills related to that setting. This form could exhibit one section of items common to all areas of field- or laboratory-based training in the Air Force and another section unique to the individual Wing, specialty or course. To accomplish this, a field- or laboratory-based counterpart to Form 281 could be developed in two or three specialty areas, thereby beginning the development of a core of items common to all specialties. These instruments could then be used as prototypes for other

Wings, specialties or courses in constructing their own sets of unique items, where such a form may be needed. One approach for transferring the needed instrument development skills to other specialties would be to conduct a brief workshop for a master instructor or supervisor from each specialty area interested in developing a field-based evaluation form. The workshop would instruct participants on the set of core items already developed and how to phrase, scale and score items consistent with the objectives of their own specialty area. The workshop would constitute a brief but practical hands-on introduction to the construction and validation of a simple field- or laboratory-based observation instrument emphasizing feedback to the instructor. Some of the content presented in Section V of this report could be relevant for such instruction.

Approximately 100 man-days are estimated for completing this recommendation, exclusive of pilot-testing.

Instructor Peer Evaluation

Peer evaluations were regularly conducted in approximately 25% of the technical courses observed by the contractor. In each instance in which peer evaluations consistently occurred, they were rated by instructors as the most valuable source of feedback they received--more valuable than ATC Form 281 and the post-observation debriefing. As noted earlier in this report, peer evaluations, although consistently used in some instructional environments, are often conducted informally at the invitation of a particular instructor, sometimes using ATC Form 281 as a general guide or starting point for the observation. However, these peer evaluators tended to supplement items on ATC Form 281 with additional indicators of teaching effectiveness that tended to "personalize" ATC Form 281 down to the specific specialty being trained for, by looking for how a particular instructor handled complex or problematic portions of specific areas of the curriculum. The concept of peer evaluation among those who used it appeared to be well accepted and valued as a meaningful source of instructional feedback. Contrary to instructor supervisors who already may be overburdened with the required administration of ATC Form 281, peers may be a possible source of additional feedback which could supplement that provided by Form 281. Since ATC Form 281 is seen by some instructors and used by some instructor supervisors as primarily a tool for evaluation not feedback, the following recommendation is made:

Section II of this paper identified 28 distinct indicators of teaching effectiveness under the more general teaching dimensions of clarity, variety, task orientation, student engagement in the learning process, and success rate. These indicators, along with their behavioral descriptors (provided in Tables 1, 2, 3, 6, 7), could be placed in the form of an observation instrument suitable for use by peers. In Section V of this paper a review of possible measurement formats indicated that the Checklist and Likert Scales represented the most practical formats for providing feedback to the ATC instructor when the criteria of acceptability to the user, cost, objectivity, increased feedback provided, ease of implementation, validity, and reliability were considered. An observation instrument suitable for peer use using a Checklist, or combination of Checklist and Likert Scale, and incorporating the teaching effectiveness indicators derived from the research literature, as described in Section II, could constitute a new feedback initiative. Instructors could be required to complete and document a specific number of peer evaluations each year, regardless of instructor status, the results of which could be kept confidential, if so desired.

Approximately 80 man-days are estimated for completing this recommendation.

Student Evaluation of Instructors

ATC Form 736 is commonly used by either individual students or a group of students to comment on an instructor's performance during a course. As was noted earlier, few individual student critiques are actually completed, although some variation across courses and instructors was

noted. Also, some group or class critiques provide only general, congratulatory or problem-specific comments that are difficult to trace back to specific indices of teaching effectiveness, thereby limiting the diagnostic value of the data. Although the open-ended response format provided by ATC Form 736 encourages originality and spontaneity, the exclusive use of such a format may also direct student responses away from more structured indices of teaching effectiveness, such as those identified in Section II of this paper. Therefore, the following recommendation is made:

In addition to ATC Form 736, a Likert-type 4- or 5-point scale could be developed to reflect the same 28 indices of teaching effectiveness (identified in Section II) and earlier recommended for use in a peer evaluation form. In this manner, student responses could complement and extend the feedback provided on the peer instructor evaluation form, since both forms would contain the same or similar indices of teaching effectiveness. Since the peer evaluation form was proposed for feedback to the instructor, the addition of a matching structured student response form could be a useful complement to it. Such an instrument could be made machine-scorable and require less than 20 minutes of class time to complete. If a Likert format is to be used, the number of student responses for each item could be provided in percentage and bar graph form and item averages could be compared to a normative sample of instructors from that content area, specialty or wing. Although the primary purpose of a structured student response form would be to provide feedback to the instructor on standardized indices of teaching effectiveness, it could also become part of the instructor's permanent record.

Approximately 80 man-days are estimated for completing this recommendation, exclusive of pilot-testing.

Technical Training Instructor's Course

The Technical Training Instructor's Course is a fully exportable instructional package that can be presented at any Technical Training Center. All new instructors must attend the course prior to their first teaching assignment. Normally, the course is taught in approximately a 3-week period. It has recently been revised and is now undergoing pilot-testing at Sheppard AFB.

The course teaches effective communication, the psychology of teaching and learning, Air Force instructor rules and regulations, and how to prepare and present lessons, among many other topics. Conspicuously missing from this curriculum, however, is reference to the teaching effectiveness indicators identified in Section II of this paper. Therefore, the following recommendations are made:

1. The Technical Training Instructor Course provides the logical place to introduce new instructors to the effective teaching behaviors identified in Section II. It is recommended that 1 hour of instruction for each of these five effective teaching behaviors be added to the Technical Training Instructor's Course. This would include five lectures that would (a) introduce the general behavior and its individual components, (b) provide a prescriptive rationale for why the particular behavior is important for effective teaching, and (c) illustrate some specific instructional practices by which the behavior can be satisfactorily exhibited in the classroom (e.g., Tables 1, 2, 3, 6, 7). Section II of this paper lays the foundation for these five lectures. In addition, it is recommended that approximately 3 hours of exercises and testing accompany the 5 hours of lecture to acquaint students with the application of these concepts in the real world of the classroom.

2. Although already suggested as an addition to the supervisor training course, it is recommended that a manual be prepared that provides standardized definitions and examples for the behaviors identified on ATC Form 281. This manual would be equally useful as a pedagogical tool

in the instructor training course for introducing the specific kinds of behavior to be observed by the instructor supervisor using ATC Form 281. As noted earlier, 80% of the teaching effectiveness indicators identified in Section II of this paper currently have counterparts on ATC Form 281, thereby providing a high degree of integration between this and the above recommendation.

Approximately 120 man-days are estimated for completing this recommendation, exclusive of pilot-testing.

REFERENCES

Air Training Command Regulation 52-8. (30 September 1986). Faculty and staff training development. Randolph AFB, TX.

Bloom, B., Englehart, M., Hill, W., & Krathwohl, D. (1956). Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain. New York: Longmans Green.

Brophy, J., & Evertson, C., (1976). Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.

Cole, N.S. (1987). Holmes Group initiative to create standards of entry to the teaching profession. Educational Measurement: Issues and Practices, Spring, p. 25.

Cornett, L. (1986, December). Incentive programs for teachers and administrators: How are they doing? Career Ladder Clearinghouse (pp. 1-27).

Crawford, J., Gage, N.L., Corno, L., Stayrouk, N., Mitman, A., Schunk, D., & Stallings, J. (1978). An experiment on teacher effectiveness and parent-assisted instruction in the third grade (Vols. 1-3). Stanford, CA: Center for Educational Research, Stanford University.

Cunningham, R.T. (1971). Developing question-asking skills. In J. Wiegand (Ed.), Developing teacher competencies. Englewood Cliffs, NJ: Prentice-Hall.

Ellett, C.D., & Garland, J.S. (1986). Examining teacher evaluation practices and policies: Results from a national survey of the 100 largest school districts. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Evertson, C., & Green, J. (1986). Observation as inquiry and method. In M.C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.). New York: Macmillan.

Fisher, C., Filby, N., Marlavec, R., Cahen, L., Dishaw, M., More, & Berliner, D. (1978). Teaching behaviors, academic learning time and student achievement. Final report of Phase III-B. Beginning teacher evaluation study (Tech. Rep. No. V-1). San Francisco: Far West Laboratory for Educational Research and Development.

Gage, N., & Berliner, D. (1984). Educational psychology (3rd ed.). Chicago: Rand McNally.

Gall, M., Ward, B., Berliner, D., Cahen, L., Winne, P., Eleshoff, J., & Stanton, G. (1978). Effects of questioning techniques and recitation on student learning. American Educational Research Journal, 15, 175-199.

Haney, W.M., & Reidy, E.F. (1987). Educational measurement: Issues and practices, Spring, pp. 4-6.

Hardy, R.A. (1986). Teacher and administrator testing: Status and trends. ERS Spectrum, 4, (4).

Hunkins, F. (1976). Involving students in questioning. Boston: Allyn & Bacon.

Kubiszyn, T., & Borich G. (1987). Educational testing and measurement. Glenview, IL: Scott Foresman.

Lehmann, I.J., & Phillips, S.E. (1987). A survey of state teacher-competency examination programs. Educational Measurement: Issues and Practice, pp. 14-18.

McDonald, F., & Elias, P. (1976). Executive summary report: Beginning teacher evaluation study, phase II 1974-1976. Princeton, NJ: Educational Testing Service.

National Commission on Excellence in Education. (1981). A nation at risk. Washington: National Commission of Excellence in Education.

Rosenshine, B., & Furst. (1973). The use of direct observation to study teaching. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally.

Servey, R. (1974). Teacher talk: The knack of asking questions. Belmont, CA: Fearon.

Soar, R., & Soar, R. (1973). Follow through classroom process measurement and pupil growth: Final report. Gainesville: University of Florida, Institute for Development of Human Resources. (ERIC Document Reproduction Service No. ED 106297)

Stallings, J., & Kaskowitz, D. (1974). Follow through classroom observation evaluation, 1972-1973 (SRI Project URU-7370). Stanford, CA: Stanford Research Institute.

APPENDIX A: SOME INSTRUCTOR EVALUATION FORMS USED IN PUBLIC SCHOOLS

CRITERIA AND INDICATORS SAMPLE--
Teacher Performance Evaluation

CRITERION 1

- 1A. The teacher demonstrates effective planning skills.
- 1B. Plans for Teaching.
- 1C. Plans for Instruction

- - - - -

INDICATORS:

- 1. Selects appropriate long-range goals.
- 2. Writes instructional objectives that are related to long-range goals.
- 3. Selects objectives at the correct level of difficulty to assure successful learning experiences for each student.
- 4. Includes teaching methods and procedures relevant to the objective.
- 5. Includes relevant student activities.
- 6. Utilizes both formative and summative evaluation procedures.
- 7. Plans appropriate time allotment.
- 8. Selects a variety of teaching methods and procedures along with a variety of student activities to use.
- 9. Writes appropriate instructional objectives.
- 10. Selects objectives congruent with district curriculum.
- 11. Incorporates appropriate materials and resources.
- 12. Develops appropriate instructional objectives.
- 13. Writes plans which are timely, current and readily available.
- 14. Follows district curriculum guides.

CRITERIA AND INDICATORS SAMPLE--
Teacher Performance Evaluation

CRITERION 2

- 2A. The teacher implements the lesson plan.
- 2B. Implements an instructional plan.
- 2C. Implements a lesson.
- 2D. Implements a lesson design.

- - - - -

INDICATORS:

- 1. Reviews and previews; provides the structure of learning.
- 2. States instructional objectives.
- 3. Provides input related to objectives.
- 4. Models activities congruent with topic being taught and provides guided practice to reinforce concepts.
- 5. Utilizes lesson summary techniques.
- 6. Provides independent practice activities.
- 7. Indicates positive directions for moving from one activity to the next.
- 8. Checks for understanding.
- 9. Provides appropriate lesson flow.
- 10. Uses appropriate methods.
- 11. Uses appropriate teaching behaviors.
- 12. Uses methods and teaching behaviors effectively.
- 13. Makes use of appropriate timing.
- 14. Reviews, previews, and verbalizes the purpose of the lesson.
- 15. Presents concepts in an orderly manner.
- 16. Demonstrates knowledge of facts and concepts of the specific lesson.
- 17. Presents appropriate background information.
- 18. Uses appropriate questioning techniques.

Criteria and Indicators Sample--Teacher Performance Evaluation

CRITERION 2 - Continued

19. Makes effective use of available instructional media and material.
20. Checks for students' understanding before going on with the lesson.
21. Provides for guided practice (where appropriate).
22. Presents instruction at an appropriate level of difficulty.

CRITERIA AND INDICATORS SAMPLE--
Teacher Performance Evaluation

CRITERION 3

3A. The teacher motivates students.

- - - - -

INDICATORS:

1. Communicates challenging scholastic expectations to students.
2. Responds positively to students.
3. Stimulates students by choosing proper materials and techniques.
4. Gives feedback to students.
5. Uses methods to stimulate creative expression.
6. Stimulates creative thinking.
7. Promotes active participation during the lesson.
8. Communicates expectations to students.
9. Moves about the classroom to facilitate learning.
10. Provides opportunities for students to experience success.
11. Stimulates student interest by utilizing appropriate content.
12. Provides variety in methods and materials throughout the year.
13. Sustains student attention and responses by planning activities at the appropriate difficulty level.
14. Provides prompt feedback to students.
15. Shows enthusiasm through teaching methods and behaviors.

HIGH ACHIEVEMENT COMMUNITY SCHOOLS

Teacher's Name	Building	Grade/Subject
----------------	----------	---------------

DIRECTIONS: For each descriptor, place a check in the column that best describes the teacher's performance for that item. After assessing the teacher's performance for each descriptor, please assess performance for the criterion by placing a check in the appropriate column.

Please check "Target for Growth" for those descriptors and criteria which the teacher should focus on for professional growth.

Specific descriptive comments must accompany any assessment where performance does not meet district standards and should also supplement "Target for Growth" and other assessments.

The performance review should be used to provide the teacher with a summative performance assessment and an opportunity to respond to the summative evaluation. Please examine performance in each criterion and check the term which represents a summative evaluation of the teacher's performance.

EXPLANATION OF THE SCALE:

NOT RATED: Insufficient data for rating performance.

MUST IMPROVE: Performance jeopardizes continued employment in the district.

NEEDS IMPROVEMENT: Performance is below the high standard required.

MEETS STANDARD: Performance meets district standards.

EXEMPLARY: Performance exceeds district standards.

TARGET FOR GROWTH: Performance area(s) in which improvement is deemed most beneficial.

CRITERION: Plans for instruction
(long and short term)

1. Develops appropriate instructional objectives.
2. Includes methods, procedures, and student activities relevant to objectives.
3. Develops both formative and summative evaluation procedures.
4. Writes plans which are timely, current, and readily available.

	NOT RATED	MUST IMPROVE	NEEDS IMPROVEMENT	MEETS STANDARD	EXEMPLARY	TARGET FOR GROWTH
1.	—	—	—	—	—	—
2.	—	—	—	—	—	—
3.	—	—	—	—	—	—
4.	—	—	—	—	—	—

CRITERION: Plans for instruction
(long and short term)

	NOT RATED	MUST IMPROVE	NEEDS IMPROVEMENT	MEETS STANDARD	EXEMPLARY	TARGET FOR GROWTH
1.	—	—	—	—	—	—

COMMENTS:

CRITERION: Fulfills professional responsibilities.

CRITERION: Fulfills professional responsibilities.	DOES NOT MEET DISTRICT STANDARD		MEETS DISTRICT STANDARD	TARGET FOR GROWTH
	_____	_____		
1. Has regular attendance and is punctual.	_____	_____	_____	_____
2. Provides and maintains accurate records.	_____	_____	_____	_____
3. Assumes, completes duties promptly and effectively.	_____	_____	_____	_____
4. Handles confidential matters or information in a professional manner.	_____	_____	_____	_____
5. Maintains good grooming.	_____	_____	_____	_____
6. Assumes school related student management responsibilities outside of classroom.	_____	_____	_____	_____
7. Works with and through organization ladder when solving problems or seeking change.	_____	_____	_____	_____
8. Adheres to school regulations and board policies.	_____	_____	_____	_____

CRITERION: Fulfills professional responsibilities.

COMMENTS:

Comments:

Plan of Action/Improvements:

Goals/Activities:

Administrator

Teacher

OBSERVATION RECORD FORM
COVINGTON PUBLIC SCHOOLS

Teacher Observed: _____ Date: _____

Observer: _____ Class: _____

I. Planning for Instruction:

- Component 1.1 Describes appropriate student objectives and teaching procedures to be used for each objective.
- Component 1.2 Develops/selects procedures or materials that provide students with appropriate practice of application on objectives.
- Component 1.3 Describes procedures and materials for assessing learner progress on lesson objectives.
- Component 1.4 Develops and maintains a record of individual student progress.

COMMENT _____

II. Implementing Instruction:

- Component 2.1 Uses techniques to develop student interest and involvement in lesson.
- Component 2.2 Uses procedures and materials which accommodate differences in the learning needs of students.
- Component 2.3 Uses instructional activities which provide students with the opportunity to be actively involved in the lesson.
- Component 2.4 Provides students with the opportunity for appropriate practice or application on objectives (including homework).
- Component 2.5 Monitors student progress or performance during lesson to identify the need for clarification, assistance, feedback or lesson modification.
- Component 2.6 Provides an orderly and attractive learning environment.

COMMENT _____

III. Managing the Instructional Environment:

- Component 3.1 Communicates expectations for student behavior.
- Component 3.2 Manages routine tasks efficiently.
- Component 3.3 Uses appropriate techniques to involve students who are off task.
- Component 3.4 Takes appropriate actions when disruptions interfere with learning of other students.
- Component 3.5 Provides assistance to students having difficulty.
- Component 3.6 Reinforces and encourages the efforts of students to maintain involvement.
- Component 3.7 Is courteous and pleasant with students.
- Component 3.8 Uses instructional time efficiently.

COMMENT _____

IV. Communicating with Students:

- Component 4.1 Communicates instructional plan, lesson content and activities in a logical sequence.
- Component 4.2 Uses effective and appropriate vocabulary and sentence structure during oral instruction.
- Component 4.3 Communicates an accurate command of subject matter, clarifying information, explanations and directions.
- Component 4.4 Uses effective written communication.
- Component 4.5 Encourages student expression of ideas, feelings, and opinions.
- Component 4.6 Communicates enthusiasm and high expectations for teaching and learning.

COMMENT _____

*** (This form is to be used to record teaching performances observed during a particular classroom visit. It is not a complete evaluation.)*

(Specimen 4)

MISSISSIPPI TEACHING COMPETENCIES AND INDICATORS

COMPETENCY I: PLANS INSTRUCTION TO ACHIEVE SELECTED OBJECTIVES.

Indicator 1. Specifies or selects learner objectives for lessons.

Indicator 2. Specifies or selects teaching procedures for lessons.

Indicator 3. Specifies or selects content, materials, and media for lessons.

Indicator 4. Specifies or selects materials and procedures for assessing learner progress on the objectives.

Indicator 5. Plans instruction at a variety of levels.

COMPETENCY II: ORGANIZES INSTRUCTION TO TAKE INTO ACCOUNT INDIVIDUAL DIFFERENCES AMONG LEARNERS.

Indicator 6. Organizes instruction to take into account differences among learners in their capabilities.

Indicator 7. Organizes instruction to take into account differences among learners in their learning styles.

Indicator 8. Organizes instruction to take into account differences among learners in their rates of learning.

COMPETENCY III: OBTAINS AND USES INFORMATION ABOUT THE NEEDS AND PROGRESS OF INDIVIDUAL LEARNERS.

Indicator 9. Uses teacher-made or teacher-selected evaluation materials or procedures to obtain information about learner progress.

Indicator 10. Communicates with individual learners about their needs and progress.

COMPETENCY IV: OBTAINS AND USES INFORMATION ABOUT THE EFFECTIVENESS OF INSTRUCTION TO REVISE IT WHEN NECESSARY.

Indicator 11. Obtains information on the effectiveness of instruction.

Indicator 12. Revises instruction as needed using evaluation results and observation data.

COMPETENCY V: USES INSTRUCTIONAL TECHNIQUES, METHODS, AND MEDIA RELATED TO THE OBJECTIVES.

Indicator 13. Uses teaching methods appropriate for objectives, learners and environment.

Indicator 14. Uses instructional equipment and other instructional aids.

Indicator 15. Uses instructional materials that provide learners with appropriate practice on objectives.

COMPETENCY VI: COMMUNICATES WITH LEARNERS. -

Indicator 16. Gives directions and explanations related to lesson content.

Indicator 17. Clarifies directions and explanations when learners misunderstand lesson content.

Indicator 18. Uses responses and questions from learners in teaching.

Indicator 19. Provides feedback to learners throughout the lesson.

Indicator 20. Uses acceptable written and oral expression with learners.

COMPETENCY VII: DEMONSTRATES A REPERTOIRE OF TEACHING METHODS.

Indicator 21. Implements learning activities in a logical sequence.

Indicator 22. Demonstrates ability to conduct lessons using a variety of teaching methods.

Indicator 23. Demonstrates ability to work with individuals, small groups, and large groups.

COMPETENCY VIII: REINFORCES AND ENCOURAGES LEARNER INVOLVEMENT IN INSTRUCTION.

Indicator 24. Uses procedures which get learners initially involved in lessons.

Indicator 25. Provides learners with opportunities for participating.

Indicator 26. Maintains learner involvement in lessons.

Indicator 27. Reinforces and encourages the efforts of learners to maintain involvement.

COMPETENCY IX: DEMONSTRATES AN UNDERSTANDING OF THE SCHOOL SUBJECT BEING TAUGHT AND DEMONSTRATES ITS RELEVANCE.

Indicator 28. Helps learners recognize the purpose and importance of topics or activities.

Indicator 29. Demonstrates knowledge in the subject area.

COMPETENCY X: ORGANIZES TIME, SPACE, MATERIALS, AND EQUIPMENT FOR INSTRUCTION.

Indicator 30. Attends to routine tasks.

Indicator 31. Uses instructional time effectively.

Indicator 32. Provides a learning environment that is attractive and orderly.

COMPETENCY XI: DEMONSTRATES HIGH EXPECTATIONS FOR LEARNERS' ACADEMIC PERFORMANCE.

Indicator 35. Conveys the impression of knowing what to do and how to do it.

Indicator 28. Helps learners recognize the purpose and importance of topics or activities.

Indicator 27. Reinforces and encourages the efforts of learners to maintain involvement.

Indicator 29. Provides feedback to learners throughout the lesson.

Indicator 31. Uses instructional time efficiently.

COMPETENCY XII: DEMONSTRATES ENTHUSIASM FOR TEACHING AND LEARNING AND THE SUBJECT BEING TAUGHT.

Indicator 33. Communicates personal enthusiasm.

Indicator 34. Stimulates learner interest.

Indicator 35. Conveys the impression of knowing what to do and how to do it.

COMPETENCY XIII: HELPS LEARNERS DEVELOP POSITIVE SELF-CONCEPTS.

Indicator 36. Demonstrates warmth and friendliness.

Indicator 37. Demonstrates sensitivity to the needs and feelings of learners.

Indicator 38. Demonstrates patience, empathy, and understanding.

COMPETENCY XIV: MANAGES CLASSROOM INTERACTIONS.

Indicator 39. Provides feedback to learners about their behavior.

Indicator 40. Promotes comfortable interpersonal relationships.

Indicator 41. Maintains appropriate classroom behavior.

Indicator 42. Manages disruptive behavior among learners.

INDICATOR 24. Uses procedures which get learners initially involved in lessons.

Comment: The indicator focuses on the procedures used by the teacher to prepare, to interest, and to motivate learners to begin the lesson.

Scale of Descriptors	Key Points in Descriptors
<p>Rating</p> <p>1. None of the descriptors is evident.</p> <p>2. One of the descriptors is evident.</p> <p>3. Two of the descriptors are evident.</p> <p>4. Three of the descriptors are evident.</p> <p>5. Four of the descriptors are evident.</p>	<p>a. The teacher helps learners recall knowledge or experiences <u>related</u> to lesson activities.</p> <p>b. Current interests of learners are the basis for new activities.</p> <p>c. A specific technique such as a discrepant event or thought-provoking questions is used by the teacher to stimulate new interest.</p> <p>d. The usefulness of activities is communicated to learners.</p> <p><u>A discrepant event is an unusual or unexpected happening.</u></p> <p>Descriptors</p> <p>a. Helps learners recall past experiences or knowledge.</p> <p>b. Uses existing interests of learners as a link to new activities.</p> <p>c. Stimulates new interests in activities with techniques such as discrepant events or thought-provoking questions.</p> <p>d. Helps learners understand what they may achieve by participating in the activities.</p>

INDICATOR 25. Provides learners with opportunities for participating.

Comment: The focus of the indicator is on the provision of opportunities for active student involvement. Both the kind of participation and amount of participation are to be considered. The level of the student must be considered in rating this indicator, for what is passive in one situation may be active for a learner on another level. The way an activity, such as a film, is conducted may often distinguish active from passive.

Scale of Descriptors

Scale of Descriptors	Key Points in Descriptors
1. Class activities require passive commitment.	1. Student activity is minimal.
2. The class is organized so that only a few learners participate actively; most appear to be bystanders.	2-4. These three descriptors differ in the number of learners given an opportunity to participate actively; they range from few to most to all. A student need not pursue an opportunity for a teacher to be rated highly.
3. Most learners have opportunity for active participation at some time in the class (e.g., small group discussion, physical manipulation of materials, physical movement, individual library work, etc.).	
4. All learners have opportunity for active participation in some type of activity.	
5. All learners have opportunity for active participation in two or more activities.	

OBSERVATION RECORD
 EVALUATION RECORD Date 10-22-86
 Date

School District NCo.-Dist. No 255-901Teacher Z Grade 12 School Year 86-87Assignment/Grade 7 School MAppraiser 1 or 2 Name of Appraiser Appraiser K Teacher Supervisor or Other Appraiser
 Circle OneObservation Date 10-22-86 Beginning Time 10:55 Ending Time 11:50Scheduled Unscheduled Subject Area(s) Observed Texas History

Directions: Circle the number(s) for each indicator. The column will be marked for each indicator. Column "A/B/E" is used to note "Absent/Below Expectation" behaviors and has a credit value of 0. Column "SE" represents the "Standard Expectation" for each indicator and has a credit value of 1. Column "EQ" represents "Exceptional Quality" for each indicator and has a credit value of 1. To give credit to Column "EQ," credit must be given in Column "SE." Exceptional Quality credit is not available to some indicators. For these, a hyphen has been placed in Column "EQ." Provide documentation when marking either Column "A/B/E" or "SE." The teacher supervisor will compute credits at the end of the appraisal period. The second appraiser will compute credits within 7 days of the formal observation. Corrections should be initiated by person filling out the form.

Instructional Strategies

1. Provides opportunities for students to participate actively and successfully.

	Columns		
	A/B/E	SE	EQ
a. varies activities	0	1	
b. interacts with students	0	1	
c. solicits participation	0	1	
d. extends	0	1	
e. provides time	0	1	
f. implements at appropriate level of difficulty	0	1	

2. Evaluates and provides feedback on student progress during instruction.

a. communicates expectations	0	1
b. monitors	0	1
c. solicits responses for assessment	0	1
d. reinforces	0	1
e. provides corrective feedback	0	1
f. reteaches	0	1

FOR EVALUATION RECORD
 DOMAIN CREDIT TOTAL
 SE credits + EQ credits

12

II. Classroom Management and Organization

Columns

3 Organizes materials and students

- a secures student attention
- b uses procedures/routines
- c gives administrative directions
- d uses seating/grouping
- e has materials/areas/facilities ready

1190

4. Maximizes amount of time available for instruction.

A graphic element consisting of a stylized dollar sign (\$) and a triangle symbol, both rendered in a dark, high-contrast style, set against a background of several horizontal lines.

- a. begins/ends
- b. implements sequence of activities
- c. maintains pace
- d. maintains focus
- e. keeps students engaged

1661

5. Manages student behavior.

- a. specifies expectations
- b. prevents off-task behavior
- c. redirects off-task behavior
- d. stops inappropriate behavior
- e. stops disruptive behavior
- f. applies rules
- g. reinforces appropriate behaviors

1
1
1
1
1
1

FOR EVALUATION RECORD
• DOMAIN CREDIT TOTAL
(SE credits + EQ credits)

17

III. Presentation of Subject Matter

6. Teaches for cognitive, affective, and/or psychomotor learning and transfer.

C & g. Teacher repeatedly reminded students that the observation skills being practiced would help them on the field trip to the pioneer village next week. "Notice color and shape of the dishes in your home. We will see examples of Indian dishes next week."

- a. begins with introduction
- b. uses content sequence
- c. relates prior/future learning
- 5 d. defines/describes
- e. elaborates critical attributes
- f. stresses generalization/principle/rule
- g. transfers
- h. closes instruction

h. During the closing reminded students to look for examples of pottery at home and in magazines.

III. Presentation of Subject Matter (continued)

Columns
A/SE SE EQ

7 Presents information accurately and clearly

a. Teacher says that Indians' only contribution to American culture is pottery. "The Karankawa Indians were potters." (They were not.) This was a lesson on pottery as a cultural contribution.

d. The teacher repeatedly stressed key points that distinguished among different Indian tribes.

8. Uses acceptable communication skills.

a. Four errors of subject-verb agreement. Examples: "There is three questions on the page." "Homework problems is for you to complete."

d. Five words misspelled on the chalkboard - environment, principle, pottery, butch, ancestor.

- a. makes no significant errors
- b. uses appropriate vocabulary
- c. explains clearly
- d. stresses points/dimensions
- e. clarifies misunderstanding

0 1
0 0
0 0
0 1
0 0

- a. uses correct grammar
- b. pronounces correctly/clearly
- c. uses accurate language
- d. demonstrates written skills

0 1
0 0
0 1
0 0

FOR EVALUATION RECORD
DOMAIN CREDIT TOTAL
(SE credits + EQ credits)

18

... Learning Environment

9. Uses strategies to motivate students for learning.

- a. relates to interests
- b. emphasizes value/importance
- c. reinforces efforts
- d. challenges

0 1
0 1
0 1
0 1

10. Maintains supportive environment.

a. b. "That's the sorriest excuse for a paper."

d. All reinforcements were generally stated - "O.K." "Fine" and accompanied each student statement.

e. Very businesslike - never called students by first name - tells rather than seeks cooperation - speaks down to students in angry tone of voice.

- a. avoids sarcasm/negative criticism
- b. maintains courteous climate
- c. encourages
- d. praises
- e. establishes rapport

0 1
0 1
0 1
0 1
0 1

FOR EVALUATION RECORD
DOMAIN CREDIT TOTAL
(SE credits + EQ credits)

5

V. Growth and Responsibilities

Columns		
A/BE	SE	EC

11. Plans for and engages in professional development.

- a. grows professionally
- b. stays current-content
- c. stays current-methods

C	1
O	0
C	0

12. Interacts and communicates with parents.

- a. initiates communications
- b. conducts conferences
- c. reports progress
- d. maintains confidentiality

O	0
O	0
O	0
O	0

13. Complies with policies, operating procedures, and requirements.

- a. follows TEA requirements
- b. follows district/campus policies/procedures
- c. performs assigned duties
- d. follows promotion procedures

O	0
O	0
O	0
O	0

14. Promotes and evaluates student growth.

- a. participates in goal-setting
- b. plans instruction
- c. documents progress
- d. maintains records
- e. reports progress

O	1
O	0
O	0
O	0
O	0

FOR EVALUATION RECORD
DOMAIN CREDIT TOTAL
(SE credits + EO credits)

16

Comments:

See me for a conference RE: Learning Environment

(The signature of the teacher indicates that he/she has reviewed and received a copy of this record.)

Teacher Z 10-28-86

Teacher Signature/Date Received

Original Copy—Central Office
Copy #2—Teacher's Supervisor
Copy #3—Teacher

Appraiser K 10-23-86

Appraiser Signature/Date Completed

11-3-87
Date of Conference
(if any)

School District NCo-Dist No 255-901Teacher Z12

132

Assignment/Grade 7thSchool M

School Year 1986-87

Domain SubtotalTeacher's Supervisor: names: T. Supervisor S

Instructions: (1) Transfer the DOMAIN CREDIT TOTALS from the Evaluation Record. Multiply Appraisal 2 by 60 to obtain the Domain Subtotal.

Domain

- I Instructional Strategies
- II Classroom Management and Organization
- III Presentation of Subject Matter
- IV Learning Environment
- V Growth and Responsibilities

Appraisal 1

<u>14</u>
<u>13</u>
<u>15</u>
<u>11</u>
<u>16</u>

Appraisal 2

<u>16</u>
<u>24</u>
<u>24</u>
<u>17</u>
<u>20</u>

Domain Subtotal

<u>60 =</u>	<u>9.6</u>
<u>60 =</u>	<u>14.4</u>
<u>60 =</u>	<u>14.4</u>
<u>60 =</u>	<u>10.2</u>
<u>60 =</u>	<u>12.0</u>

Other Appraiser: (names) Appraiser K

Instructions: (1) Transfer the DOMAIN CREDIT TOTALS from the Evaluation Records. Multiply Appraisal 2 by 40 to obtain the Domain Subtotal.

Domain

- I Instructional Strategies
- II Classroom Management and Organization
- III Presentation of Subject Matter
- IV Learning Environment
- V Growth and Responsibilities

Appraisal 1

<u>12</u>
<u>17</u>
<u>17</u>
<u>9</u>
<u>16</u>

Appraisal 2

<u>12</u>
<u>17</u>
<u>17</u>
<u>9</u>
<u>16</u>

Domain Subtotal

<u>40 =</u>	<u>4.8</u>
<u>40 =</u>	<u>6.8</u>
<u>40 =</u>	<u>6.8</u>
<u>40 =</u>	<u>3.6</u>
<u>40 =</u>	<u>6.4</u>

Domain Scores:

(1) Compute the Domain Total by adding the domain subtotals of both Appraisers.

(2) Consult the Score Conversion Chart to convert each Domain Total to a Domain Score.

	Teacher Supervisor	Other Appraiser	Domain Subtotal	Domain Total	Domain Score
I	<u>9.6</u>	<u>4.8</u>		<u>14.4</u>	<u>4.3</u>
II	<u>14.4</u>	<u>8</u>		<u>21.2</u>	<u>4.9</u>
III	<u>14.4</u>	<u>8</u>		<u>21.2</u>	<u>4.5</u>
IV	<u>10.2</u>	<u>8</u>		<u>13.8</u>	<u>4.9</u>
V	<u>12.0</u>	<u>6.4</u>		<u>18.4</u>	<u>4.9</u>

Overall Summary Performance Score:

(1) Enter the Domain Scores in the spaces provided below and compute the sum.

(2) Divide the total by 5.

Domain Scores for:

Domain I	Domain II	Domain III	Domain IV	Domain V	Average
<u>4.3</u>	<u>4.9</u>	<u>4.5</u>	<u>4.9</u>	<u>4.9</u>	<u>23.5</u> - 5 - <u>4.7</u>

(3) Record the Overall Summary Performance Score as an average of the domain scores except that the overall performance score can be no higher than

(a) 1.9 if one or more of the domain scores are 1 to 1.9 inclusive, or

(b) 2.9 if two or more of the domain scores are 2 to 2.9 inclusive and no score is 1 to 1.9.

Overall Summary Performance Score =

4.7

(The signature of the teacher indicates that he/she has received a copy of the Appraisal Record and the results of the performance appraisals have been discussed at a conference.)

Teacher Z

Teacher's Signature

Signature of other appraiser(s) if present at conference

4-17-87

Date of Conference

APPENDIX B: ATC CLASSROOMS VISITED BY THE CONTRACTOR

ATC Courses of Instruction Observed by the Contractor:

Lackland AFB

- Recruiter training course
- Social actions course
- Security police specialist course

Keesler AFB

- Air traffic control operator course
- 43E I/O Radar maintenance course
- Automatic radar tracking specialist course
- Basic electronics course

Sheppard AFB

- Dental assistant specialist course
- Advanced X-ray system course
- Diet therapy specialist course
- Medical service specialist course
- Technical training instructor course
- Electrician specialist course
- Telephone specialist course

Lowry AFB

- Inventory management specialist course
- B52 tail gun avionics course
- F15 Aircraft armament systems specialist course

APPENDIX C: ATC INSTRUCTOR EVALUATION FORMS

INSTRUCTOR EVALUATION CHECKLIST

DATE

INSTRUCTIONS: Check (✓) each item in Section I as: O - Outstanding; S - Satisfactory; NI - Needs Improvement. Items which are not applicable to the type of lesson presented will be (NA). Ratings of "Needs Improvement" and "Outstanding" will require written justification in Section I. Ratings of "Needs Improvement" will require a follow-up evaluation in Section IV within 30 days.

NAME OF INSTRUCTOR (Last, First, Middle Initial)	GRADE	ORGANIZATION	TIME STARTED	TIME COMPLETED
NAME OF EVALUATOR (Last, First, Middle Initial)	GRADE	COURSE	SUBJECT	

SECTION I.	ITEMS	RATING (/)				COMMENTS
		O	S	NI	NA	
A. PREPARATION						
1. Lesson plan current, personalized, and appropriate. Lesson plan signed by instructor's supervisor.						
2. Classroom neat and orderly. Seating arrangement appropriate. Items identified needing repair, such as burned out lights.						
3. Training materials, aids, and equipment available prior to the start of class.						
4. Necessary materials distributed effectively.						
B. CONTENT/STRUCTURE OF PRESENTATION						
1. Objectives clearly stated and sequence for lesson briefly outlined.						
2. All objectives covered. Lesson flowed smoothly from point to point.						
3. Main points reviewed in conclusion (or internal summary given if lesson not ended.)						
C. PRESENTATION SKILLS						
1. Eye contact made with students.						
2. Movement and gestures natural and appropriate, not forced or distracting.						
3. Instructor poised, enthusiastic, and confident.						
D. COMMUNICATION SKILLS						
1. Correct enunciation and grammar used.						
2. Excessive use of distracting mannerisms such as "Ahs" and "OKs" not evident.						
3. Instructor's voice quality, volume, and variation (pitch, rate, inflection) appropriate for group and classroom size.						
E. DEMONSTRATION SKILLS						
1. Skills properly introduced and demonstrated.						
2. Students involved in demonstration, if appropriate.						
F. USE OF TRAINING AIDS/MATERIALS						
1. Training aids/instructional materials/equipment listed in POI used by the instructor/students during lesson.						
2. Chalkboard and other visual aids used in an effective manner.						

SECTION I. (Continued) ITEMS	O	S	NI	NA	COMMENTS
G. QUESTION/ANSWER TECHNIQUES					
1. Questions phrased clearly and to the point.					
2. Questions appropriate for the lesson.			-		
3. Variety of question types used. Types of questions used adjusted to the situation.					
4. Student questions answered adequately.					
H. MANAGEMENT					
1. Proper control of class maintained.					
2. Appropriate techniques used to assist and motivate students.					
3. Time managed appropriately. Lesson well paced.					
4. Instructor properly used the MIR instructor.					
5. Student participation encouraged.					
I. SUPERVISING GROUP ACTIVITY SKILLS					
1. Clear instructions provided to the group.					
2. Group members encouraged to participate.					
J. MEASUREMENT					
1. Progress check and performance/written tests administered properly.					
2. Proper critique procedures followed.					
K. PERSONAL QUALITIES					
1. Thorough knowledge of subject matter demonstrated.					
2. Instructor's professionalism set the proper example for bearing, behavior and dress.					
3. Positive rapport with students established.					
L. OTHER					
1. Students in compliance with appropriate dress and appearance standards.					
2. Importance of safety emphasized and compliance with safety standards ensured.					
3.					
4.					
5.					
6.					
7.					

SECTION II.**OVERALL PERFORMANCE RATING**

CHECK THE OVERALL PERFORMANCE AS DERIVED FROM THE EVALUATION IN SECTION I. WHEN COMPUTING THE OVERALL PERFORMANCE RATING, ITEMS MARKED "NOT APPLICABLE" ARE NOT COUNTED.

 OUTSTANDING

- More than 75% of the items rated received "Outstanding". No items were rated "Needs Improvement".
- Majority of the items rated received "Satisfactory" or "Outstanding". Less than 25% of the items rated received a rating of "Needs Improvement".
- More than 25% but less than 50% of the items rated received a rating of "Needs Improvement".
- More than 50% of the items rated received a rating of "Needs Improvement".

 SATISFACTORY

- Majority of the items rated received "Satisfactory" or "Outstanding". Less than 25% of the items rated received a rating of "Needs Improvement".

 NEEDS IMPROVEMENT

- More than 25% but less than 50% of the items rated received a rating of "Needs Improvement".

 UNACCEPTABLE

- More than 50% of the items rated received a rating of "Needs Improvement".

SECTION III.**ACKNOWLEDGEMENT OF INTERVIEW****SIGNATURE OF EVALUATOR****DATE AND SIGNATURE OF INSTRUCTOR EVALUATED****DATE, SIGNATURE AND GRADE OF IMMEDIATE SUPERVISOR OF EVALUATOR****SECTION IV.****RECOMMENDATIONS FOR IMPROVEMENT/GENERAL COMMENTS****SECTION V.****FOLLOW-UP EVALUATION (Within 30 days)****SIGNATURE OF EVALUATOR****DATE AND SIGNATURE OF INSTRUCTOR EVALUATED**

STUDENT CRITIQUE

INSTRUCTIONS: Use this form to evaluate training, student group or base support facilities and services. Use a separate form for each area being evaluated; i.e., do not comment on training and base support facilities on the same form.

COURSE NO	COURSE TITLE		PERIOD OF TRAINING		
			-	FROM	TO
NAME OF STUDENT (Optional)		GRADE	DATE	CLASS NO	SHIFT
					SQUADRON
CRITIQUE ON (Check one)			TYPE OF CRITIQUE		
<input type="checkbox"/> TRAINING <input type="checkbox"/> SCHOOL SQ <input type="checkbox"/> BASE SUPPORT FACILITIES			<input type="checkbox"/> INDIVIDUAL <input type="checkbox"/> GROUP <input type="checkbox"/> ORAL-TYPE		
CRITIQUE RESULTS REQUESTED		ADDRESS/PHONE NUMBER (For later contact, if critique results are requested)			
<input type="checkbox"/> YES <input type="checkbox"/> NO					

STUDENT COMMENTS (Tell who, what, when, where, why and how) (Continue on reverse if necessary)

FOLLOW-UP ACTION TAKEN

TIME	DATE	METHOD OF STUDENT CONTACT		
------	------	---------------------------	--	--

APPENDIX D: ARMY AND NAVY INSTRUCTOR EVALUATION FORMS

USACMLS INSTRUCTOR ASSESSMENT

Instructor _____ Div _____ Dir _____ ITC Qualified Yes
 Lesson No. _____ Title _____
 Course _____ Class Location _____ Date & Time _____ / _____
 Evaluator _____ ITC Qualified Yes No Assessment Form Qualified Yes

COMPUTATIONS

<u>SECT</u>	<u>EARNED POINT SUBTOTALS</u>	<u>NONRATED FACTORS</u>
A	_____	_____
B	_____	_____
C	_____	_____
D	_____	_____
E	_____	_____
TOTAL	(f)	(g)

Scoring base 600
 Minus total nonrated factors (g) $\times 6 =$ (h)
 Aligned base score (i)
 $\times 100 =$ **SCORE*** (k)

* Round score to nearest whole number.

SECTION A. Implementation/Statement of Objectives:

1. Gains attention/motivates at beginning of class.
2. Identifies objective(s) at beginning of class.
3. Tell students application of objective(s) and consequence(s) of non-performance of tasks(s); ties lesson in to past instruction.

Comments: _____

<u>FACTOR WEIGHT</u>	<u>EVALUATION</u>	<u>EARNED POINTS</u>
5	NR 1 2 3 4 5 6	_____
5	NR 1 2 3 4 5 6	_____
4	NR 1 2 3 4 5 6	_____
<u>14</u>	<u>Subtotal</u>	_____

SECTION B. Lesson Presentation (Body)

1. Lesson plan well structured/current/LAW policy letter.
2. Trains LAW lesson plan.
3. Presents lesson plan in an interesting and innovative way (stories, humor, examples).
4. Uses training aids professionally.
5. Uses transitions and summaries effectively.
6. Concludes and ties lesson together and uses an appropriate closing statement.

Comments: _____

6	NR 1 2 3 4 5 6	_____
8	NR 1 2 3 4 5 6	_____
6	NR 1 2 3 4 5 6	_____
5	NR 1 2 3 4 5 6	_____
4	NR 1 2 3 4 5 6	_____
5	NR 1 2 3 4 5 6	_____
<u>34</u>	<u>Subtotal</u>	_____

SECTION C. Instructor Communication Skills

<u>FACTOR</u>	<u>WEIGHT X</u>	<u>EVALUATION</u>	<u>EARNED POINTS</u>
1. Phrases questions to elicit student thought/ correct answers.	4	NR 1 2 3 4 5 6	_____
2. Uses proper questioning techniques.	4	NR 1 2 3 4 5 6	_____
3. Speech & Voice.	5	NR 1 2 3 4 5 6	_____
4. Gestures.	3	NR 1 2 3 4 5 6	_____
5. Eye contact.	3	NR 1 2 3 4 5 6	_____
6. Movement.	1	NR 1 2 3 4 5 6	_____
	<u>20</u>	Subtotal	_____

Comments: _____

SECTION D. Instructor Preparation

1. Civilian dress is tasteful/Military appearance is IAW AR 670-1 and local regulation.	1	NR 1 2 3 4 5 6	_____
2. Is in control of class.	4	NR 1 2 3 4 5 6	_____
3. Appears highly motivated.	4	NR 1 2 3 4 5 6	_____
4. Is relaxed and confident in interaction with students.	3	NR 1 2 3 4 5 6	_____
5. Shows concern, understanding, and willingness to help students.	3	NR 1 2 3 4 5 6	_____
6. Exhibits competence in subject matter.	6	NR 1 2 3 4 5 6	_____
7. Ensures that doctrine trained is current and correct.	6	NR 1 2 3 4 5 6	_____
	<u>27</u>	Subtotal	_____

Comments: _____

SECTION E. Training Site Management

1. Ensures that physical arrangement facilitates instruction and student involvement.	2	NR 1 2 3 4 5 6	_____
2. Ensures that Visitor Folder is IAW USACMLS Policy Letter #67.	1	NR 1 2 3 4 5 6	_____
3. Ensures that AIs (when available) are used appropriately.	2	NR 1 2 3 4 5 6	_____
	<u>5</u>	Subtotal	_____

Comments: _____

GENERAL COMMENTS: _____

(NAVY)

INSTRUCTIONAL EFFECTIVENESS CHECKLIST

Instructions: Answer yes or no to relevant questions; do not answer irrelevant questions nor force inappropriate yes or no answers. Use the back of the form to provide additional details as required. Answer the following questions as the lesson progresses.

LEARNING ORIENTATION	Yes	No
Are students given objectives?		
Clarified/amplified (if necessary)?		
Were the students motivated in terms of "why" the content should be learned?		
Were the students motivated re "how" the content should be used?		
Are students told how they will be tested?		
Are students told what they will be tested on?		
Did instructor frequently orient students to job of learning? Was information about the instructional process and what the students were supposed to be doing/learning given?		

INSTRUCTOR BEHAVIOR	Yes	No
Was peer instruction used?		
Were external rewards given?		
Did instructor establish relationship w/ Ss by introducing her/himself? by displaying course/unit/module name? by creating interest in subject? by displaying enthusiasm? by soliciting class cooperation and involvement?		
Is the instructor's voice level and enunciation adequate?		
Was the instructor free from any distracting mannerisms (e.g. ticks, twitches)?		
Did it appear that the instructor adequately prepared for the lesson?		
Did the instructor urge students to take notes?		
At appropriate points in lecture, did the instructor pause and indicate to students they should take notes?		
Did Instructor allow students to learn by doing when possible?		
Did instructor insure that students were actively involved?		
Did instructor monitor student progress?		
Did instructor provide assistance when necessary?		

INSTRUCTOR BEHAVIOR (CON'T)	Yes	No
Was IG used properly?		
Was Curriculum Outline adhered to?		
Was student comprehension checked?		
Were proper questioning techniques used?		
Was class control maintained?		
Was eye contact maintained?		
Was summary/critique appropriate?		

MEDIA / MATERIALS	Yes	No
What types of media were used? -	x	x
Were they appropriate?		
Were the transparencies,etc, technically correct (easy to read, layout attractive, etc.)		
Were there sufficient number of instructional aids? (example, could transparencies be used instead of blackboard?)		
Were they effectively used in separating the various instructional activities/phases/summary?		
ENVIRONMENT / SAFETY		
Describe significant environmental factors involving light, background noise, temperature, general attractiveness of classroom, etc.	x	x
Was the equipment operational?		
Were necessary tools and test equipment available?		
Was lab clean and free of safety hazards?		
Was two man rule observed?		
Were safety precautions announced?		
Were students aware of emergency procedures?		
Was emergency first aid procedure posted and visible?		
Were high voltage areas clearly marked?		

STUDENT BEHAVIOR	Yes	No
Did it appear the students achieved the objectives?		
Did it appear the student took adequate amount of notes?		
How involved were the students in the learning process? Circle the most appropriate category listed below that describes the majority of the students.		
ACTIVE Student asks questions, answers INVOLVEMENT questions, volunteers extra information generates information him/herself. May go beyond information immediately given and relate to other known information.		
ATTENTIVE Student clearly understands and is tracking INVOLVEMENT what teacher is saying. Seems interested. Indicated by alert faces, heads nodding positively, alert posture, smiling, talking to self.		
PASSIVE Student seems bored and is only marginally INVOLVEMENT paying attention to what is going on in the class. Is able to answer straightforward questions, but does not exhibit a great deal of interest in what is being taught.		
NO Students' eyes may be glazed over and may INVOLVEMENT have difficulty staying awake. Can't answer question if called upon. May not even seem to know what lesson instructor is addressing.		